

GenCore version 5.1.6  
Copyright (c) 1993 - 2006 Compugen Ltd.

## OM protein - protein search, using SW model

Run on: January 20, 2006, 19:14:26 ; Search time 134 Seconds

(without alignments) 2167.383 Million cell updates/sec

Title: US-09-463-733-1  
perfect score: 3442  
Sequence: 1 MDENAIIRAAIFIQKTYRHHQ..... VEHIDIDPTCESKVDPKKS 661

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext: 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

- 1: A\_Geneseq\_21:\*
- 2: geneseqD1980s:\*
- 3: geneseqD2000s:\*
- 4: geneseqD2001s:\*
- 5: geneseqD2002s:\*
- 6: geneseqD2003s:\*
- 7: geneseqD2003bs:\*
- 8: geneseqD2005s:\*
- 9: geneseqD2005bs:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match Length	DB ID	Description
1	3315	96.3	637 4 ABB71489	Abb71489 Drosophili
2	1305	37.9	722 8 ADN22998	Adn22998 Bacterial
3	1271.5	36.9	653 4 ABA7250	Abd47250 Human
4	1271.5	36.9	653 8 ADT6673	Adt6673 Human pro
5	1164.5	33.8	613 7 ADJ95096	Adj95096 Novel
6	1164.5	33.8	613 7 ADJ95098	Adj95098 Novel
7	581	16.9	488 8 ADY05391	Ady05391 Plant ful
8	577	16.8	520 4 ABB63840	Abb63840 Drosophili
9	571	16.6	500 4 ABG0989	Abg0989 Novel hum
10	570	16.6	494 5 ABG0989	Abg0989 Novel hum
11	570	16.6	499 4 ABR9722	Abg09722 Novel cel
12	570	16.6	499 7 ADR64103	Ade64103 Human Pro
13	570	16.6	499 8 ABM82127	Abm82127 Tumour-as
14	569	499	7 ADB64101	Ade64101 Rat Prote
15	562.5	16.3	482 5 ABE6244	Aab16244 Tomato in
16	562.5	16.3	485 5 ABE6236	Aab16236 Tomato in
17	562	16.3	499 4 ABR9733	Ade09733 Protein p
18	560	16.3	499 4 AAE9731	Aae09731 Protein p
19	556	16.2	499 4 AAE09732	Aae09732 Protein p
20	519	15.1	473 8 ADR54165	Adr44165 Bacterial
21	502.5	14.6	598 8 ADN1438	Adn21438 Bacterial
22	496.5	14.9	714 4 ABG09988	Abg09988 Novel hum
23	479.5	13.9	513 7 ADK64862	Adk64862 Protein s

## ALIGNMENTS

RESULT 1	ID	ABB71489 standard; protein; 637 AA.
XX	XX	AC
XX	XX	ABB71489;
XX	XX	DT
XX	XX	26-MAR-2002 (first entry)
DB	XX	Drosophila melanogaster polypeptide SEQ ID NO 41259.
KW	XX	Drosophila; developmental biology; cell signalling; insecticide; pharmaceutical.
OS	XX	Drosophila melanogaster.
RN	XX	W0200111042-A2.
PP	XX	
PR	XX	23-MAR-2000; 2000US-019167P.
PR	XX	11-JUL-2000; 2000US-00614150.
(PBKB ) PB CORP NY.	XX	
PI	XX	Venter JC, Adams M, Li PW, Myers EW;
DR	XX	WRI; 2001-656860/75.
N-PSDB; ABL15592.	XX	
PT	XX	New isolated nucleic acid detection reagent for detecting 1000 or more genes from Drosophila and for elucidating cell signaling and cell-cell interactions.
PT	XX	Disclosure: SEQ ID NO 41259; 21pp + Sequence Listing; English.
CC	CC	The invention relates to an isolated nucleic acid detection reagent capable of detecting 1000 or more genes from Drosophila. The invention is useful in developmental biology and in elucidating cell signalling and cell-cell interactions in higher eukaryotes for the development of insecticides, therapeutics and pharmaceutical drugs. The invention discloses genomic DNA sequences (ABL16176-ABL1051), expressed DNA sequences (ABL0840-ABL1675) and the encoded proteins (ABL57737-ABL72072). The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences

XX	Sequence	637 AA;	XX	PD	18-DEC-2003.
Query Match	96.3%	Score 3315; DB 4; Length 637;	XX	PP	20-FEB-2003; 2003US-00369493.
Best Local Similarity	100.0%	Pred. No. 2. 6e-304; Mismatches 0; Indels 0; Gaps 0;	XX	PR	21-FEB-2002; 2002US-0360339P.
Matches	637;	Conservative	XX	PA	(CAOY) CAO Y.
Qy	25	MORRCNMQIPQNLKEYASQDQARLYKEPNFLIKHMPQAAGRKNQYQGSAHNTSVLDDKDL	PA	(HINK) HINKLE G J.	
Db	1	MORRCNMQIPQNLKEYASQDQARLYKEPNFLIKHMPQAAGRKNQYQGSAHNTSVLDDKDL	PA	(SLAT) SLATER S C.	
Qy	85	VZBFGDVTINAKELBLPIRKHNHIDLLIDVFRKKGRLHPKVYVALLRERAKSILQLENISP	PA	(CHEN) CHEN X.	
Db	61	VSEFGDVIINKELBLPIRKHNHIDLLIDVFRKKGRLHPKVYVALLRERAKSILQLENISP	PA	(GOLD) GOLDMAN B S.	
Qy	145	VSTAVSQDTVCGDILHGKLDLWHLNGLSSSNPYTFNGDFVDRGKGRSLFVLLS	XX	Cao Y., Hinkle GJ, Slater SC, Chen X., Goldman BS;	
Db	121	VSTAVSQDTVCGDILHGKLDLWHLNGLSSSNPYTFNGDFVDRGKGRSLFVLLS	XX	DR	
Qy	205	LYLAEPNAVPLNGNHEDSVNARYGFTREVESKYPRNHKRLADEVYRMLPLGSVN	XX	WPI; 2004-061375/06.	
Db	181	LYLAEPNAVPLNGNHEDSVNARYGFTREVESKYPRNHKRLADEVYRMLPLGSVN	XX	Cao Y., Hinkle GJ, Slater SC, Chen X., Goldman BS;	
Qy	265	SRVLIVHGGFSISTSLSLKLISDKTDRGKYSILARPLPTGEPLKTEWQOIQFDIMWSPOAT	XX	XX	
Db	241	SRVLIVHGGFSISTSLSLKLISDKTDRGKYSILARPLPTGEPLKTEWQOIQFDIMWSPOAT	XX	CC	
Qy	325	MGCYVNPNTLRGAGTMFGPDTVDPLQLRPLTQDFPLQTKTEWQOIQFDIMWSPOAT	CC	The invention relates to a recombinant DNA construct comprising a promoter functional in a plant cell, where the promoter is positioned to provide for expression of a polynucleotide encoding a polypeptide from a microbial source, useful for producing plants with improved properties.	
Db	301	MGCYVNPNTLRGAGTMFGPDTVDPLQLRPLTQDFPLQTKTEWQOIQFDIMWSPOAT	CC	Claim 1; SEQ ID NO 5651; 122pp; English.	
Qy	385	YAGSGNKGAYIRLNLQMLPHFQYIASQTRLSPKORMGVESSALKHLAVRMHD	CC	CC	
Db	361	YAGSGNKGAYIRLNLQMLPHFQYIASQTRLSPKORMGVESSALKHLAVRMHD	CC	CC	
Qy	445	ELDEDEFRKYKDPKDSGYTSISHWKCKMVENTKUGLPWLRLDKLAPGTDQKYNMRIDL	CC	CC	
Db	421	ELDEDEFRKYKDPKDSGYTSISHWKCKMVENTKUGLPWLRLDKLAPGTDQKYNMRIDL	CC	CC	
Qy	505	LDTDVILRAEADGMSVMDALYANKASLVAIFNITDAHSGETLDEETATIDLIVAHMPG	CC	CC	
Db	481	LDTDVILRAEADGMSVMDALYANKASLVAIFNITDAHSGETLDEETATIDLIVAHMPG	CC	CC	
Qy	565	AVSKAEMLKCRMMLDINGDKVULNEFLAFLPSLDRKKEQDENIRRSTPSVAKTA	CC	CC	
Db	541	AVSKAEMLKCRMMLDINGDKVULNEFLAFLPSLDRKKEQDENIRRSTPSVAKTA	CC	CC	
Qy	625	TDPVTLIADKISGNTLVWEHDIDPTDCESKVKIDPKS	CC	CC	
Db	601	TDPVTLIADKISGNTLVWEHDIDPTDCESKVKIDPKS	CC	CC	
RESULT 2	ADN22998	ADN22998 standard; protein: 722 AA.	XX	Sequence	722 AA;
Query Match	37.9%	Score 1305; DB 8; Length 722;	XX	Query Match	37.9%; Score 1305; DB 8; Length 722;
Best Local Similarity	41.6%	Pred. No. 9. 6e-114; Mismatches 201; Indels 56; Gaps 13;	XX	Best Local Similarity	41.6%; Pred. No. 9. 6e-114; Mismatches 201; Indels 56; Gaps 13;
Matches	266;	Conservative	XX	Matches	266;
Qy	4	NATRAAFIQQYRGRARREMRORCNRQIQNLKEYASQDQARLYKEPNFLIKHMPQAAGRKNQYQGSAHNTSVLDDKDL	DB	4	NATRAAFIQQYRGRARREMRORCNRQIQNLKEYASQDQARLYKEPNFLIKHMPQAAGRKNQYQGSAHNTSVLDDKDL
Db	90	STIKSATLIIOKWYRCARLEARSLARRATWQITFALSYAGEPOLKVRWWTIXDFFAVIR	DB	57	STIKSATLIIOKWYRCARLEARSLARRATWQITFALSYAGEPOLKVRWWTIXDFFAVIR
Qy	58	HMPQAAGRKNQYQ-SAHNTSVLDD-KDLDVBEFGDVIINKEL	DB	149	HMPQAAGRKNQYQ-SAHNTSVLDD-KDLDVBEFGDVIINKEL
Db	150	AMARKKKGCTGNGRSPMLALSHTYAKPSIMDSGCTEVKOMLEDPSPTVNDIDRKYKGP	DB	97	AMARKKKGCTGNGRSPMLALSHTYAKPSIMDSGCTEVKOMLEDPSPTVNDIDRKYKGP
Qy	98	--LPTRKHNHIDLLIDVFRKKGRLHPKVYVALLRERAKSILQLENISPVSTAVSQDT	DB	154	--LPTRKHNHIDLLIDVFRKKGRLHPKVYVALLRERAKSILQLENISPVSTAVSQDT
Db	210	TISLPLDKPOVAKMIAF--KUNKVLPYVLMHLRARKFKMFPMSSVSRSITISNQV	DB	267	TISLPLDKPOVAKMIAF--KUNKVLPYVLMHLRARKFKMFPMSSVSRSITISNQV
Qy	155	VCGDLAKGKGLDILWHLNGLSSSNPYTFNGDFVDRGKGRSLFVLLS	DB	214	VCGDLAKGKGLDILWHLNGLSSSNPYTFNGDFVDRGKGRSLFVLLS
Db	268	ICGDLMKGFDDCLITYKNGYPSVDPYTFNGDFVDRGKGRSLFVLLS	DB	327	ICGDLMKGFDDCLITYKNGYPSVDPYTFNGDFVDRGKGRSLFVLLS
OS	Bacteria.		OS	215	LNRRGNHEDSVNARYGFTREVESKYPRNHKRLADEVYRMLPLGSVLNSRVLVHGGF
PN	US200333675-A1.		PN	328	LNRRGNHEDSVNARYGFTREVESKYPRNHKRLADEVYRMLPLGSVLNSRVLVHGGF

OY 275 SDPSNLIDLKISIDRKGKVSLIRPLPDTGRPLDK-----TEWQIQIDMWSDDPOATMG 327  
 CC CC intracellular calcium which provides targets for screening assays that  
 DB 388 SDQEVSKLKDCKIPHRHFRFOSVLRPVKNCMSEKSAVNUDEWQMDIMWSDPKQNGC 447  
 OY 328 VPIVLRGAGWWRDVTNPLQRHLSTVRSHECKPGRHFDNKITIISASNYAI 387  
 DB 448 WPNVFRGGSYEGADITASPLERKHFRFLVRSHHECKPGRGEGEFSHNNTCLTFSASNYET 507  
 OY 388 GSNGKAYIR-LNNQLMMPHQYQVISAASOTKRLS-FKORMGIVESSALKELAVMRDRDE 445  
 DB 508 GSNGKAYIKLGSKQPRIVQYM-ASCTHRKTILREALGVWESASVKEKGKLSFH 565  
 OY 446 LEDEFRKDPKDGSYISTSHWCKMVNTKGLPWRLLRKDQLAP-GTDSQKVNRTDL 504  
 DB 566 LOKEPEIMDIKEKSCKLPIIWKSDCVERTIGLNIFWIALAKVANLSEBGKVNMYKEDRI 625  
 OY 505 LDTVILAEADGMVNDALYANKASJVAIFIIDANSGRBLFESTAIDLWAMPG 564  
 DB 626 AOVGGTHAQKD--IYESLYRHKSTLETFRMKONGQSMKFRIDACEVL----G 677  
 OY 565 AVSKAEM---LIEKCRMADMLNGDGKVLDNLNEFLAFRLSD 599  
 DB 678 KVKRPLQLDVIQSQIAESIDFNKGDIFDNLNEFLAFRLVD 717

RESULT 3

AAB47250 standard; protein; 653 AA.

ID AAB47250; XX  
 AC AAB47250; XX  
 DT 18-JUL-2001 (first entry)  
 XX Human PPT.  
 XX Mitochondrial calcium-binding protein; p95.6/yN52; calcium, drug screening; sequence extraction; PPT; calcium binding motif.  
 XX Homo sapiens.  
 PN W0200130830-A2.  
 XX 03-MAY-2001.  
 XX 26-OCT-2000; 2000WO-US029787.  
 XX PR 26-OCT-1999; 99US-00427857.  
 PR 26-OCT-1999; 99US-00427855.  
 PR 07-SEP-2000; 2000US-00657253.  
 PA (MITO-) MITOKOR.

PI Murphy AM, Fahy ED, Wiley SE, Clevenger W;  
 XX DR WRI; 2001-308616/32.  
 DR N-PSDB; AAC58504.

PT New Caenorhabditis elegans mitochondrial calcium-binding protein, identified by protein motif database searching, useful for regulating intracellular calcium ions and in drug screening assays.

Example 2; Fig 13; 17pp; English.

RESULT 4

ADT6673 standard; protein; 653 AA.

ID ADT6673; XX  
 AC ADT6673; XX  
 DT 16-DRC-2004 (first entry)  
 DE Human protein phosphatase, EF hands-1 protein.  
 XX pain regulation; screening; LuxP; PRP-19; phosphatidylinositol synthase; valosin; interleukin 6 receptor subunit beta; aspartate aminotransferase; neuronal immediate early gene; heat shock protein 27; Hsc70; calmodulin; syntaxin binding protein 1; spliceosomal protein SAP 155; neurodin 1; basanacan; leukotriene A4 hydrolase; chondromodulin 1; Mab4/phosphatidylinositol-4-phosphate 5 kinase; 26S proteasome UB P112; US Z; UE P44.5; ingeain; 3-hydroxy-3-methylglutaryl-Co A synthase; phosphoglycerate kinase; RNA polymerase II TF SII; p18 UE;

CC proteins involved in binding, transport or other regulation of intracellular calcium which provides targets for screening assays that identify agents capable of altering a mitochondrial activity

CC sequence 653 AA.

Query Match 36.9%; Score 1271.5; DB 4; Length 653;  
 Best Local Similarity 40.3%; Pred. No. 1\_2e-110; Matches 258; Conservative 129; Mismatches 204; Indels 49; Gaps 11;  
 Matches 258; Conservative 129; Mismatches 204; Indels 49; Gaps 11;

OY 3 EMAIRAMIFQKWRHARREMORRNWQIOTNLBEASBOPAELKPFENDLIKHNQQA 62  
 DB 14 DTSIRALAUITIONWYRGKVKARLKHQHALTIFQSIYADEOGCMQUSTPFESMLENTHI 73  
 OY 63 AGRKNOQGSANTHSVLDKD-DLV-BFGIVNA-KIELLIRKHNIDLUDVPRKRG 117  
 DB 74 HKERELBRUNQNSLBSBOMDRMDYVDSIDVPSYNGRQLOFFLTCTUDLLEAFKQQ- 132  
 OY 118 NRHLHPKVVALIREAAKSILQKLPNISPVSTAVSQVTVCGDLAIGKUDLUVLHKQNLPS 177  
 DB 133 -LHAKTVELEVETKULKQMNPFHTQTSFSKEVTCGDLAIGKUDLFLPFTKQNLPS 191  
 OY 178 SSNPYVNGDFYDGRKGKGLLEVLLLSSLYLAPNPAVTLANGHEDSYNARYGPIREVES 237  
 DB 192 ERNPYVNGDFYDGRKGKGLLEVLLLSSLYLAPNPAVTLANGHEDSYNARYGPIREVES 251  
 OY 238 KPRNHRILATIDEVTRWLPGSVLNSRLVHGGSDSTSIDLKSIDRKGKVSLRP 297  
 DB 252 KYKLHGRILQIQLKEPTAWLPIGTIVUNELIIVHGGSSETTOLNLARVERNMOKSFLP 311  
 OY 298 PL-----TDGERP---LDTKEMQIDMWSDDPOATMGCV 328  
 DB 312 PTETNRHDHTDSKHKQGVTFNAHGRKTKNGSPTEHTEHEWQDIDLWSPDRGKQCP 371  
 OY 329 PNTLRLGAGWWRDVTNPLQRHLSTVRSHECKPGRHFDNKITIISASNYAI 388  
 DB 372 PNTCRGCGCYGPDVTSKILNKVQKOLMKLIRSHHECKPGRHFDNKITIISASNYAI 388  
 OY 389 SNKGAYIR-LNNQLMMPHQYQVISAASOTKRLS-FKORMGIVESSALKELAVMRDRDE 445  
 DB 432 SNGKAYIKLGSKQPRIVQYM-ASCTHRKTILREALGVWESASVKEKGKLSFH 487  
 OY 446 LEDEFRKDPKDGSYISTSHWCKMVNTKGLPWRLLRKDQLAP-GTDSQKVNRTDL 505  
 DB 488 LTRAFQLODHRKSKLQVSQWAFCMENIILGNNLPHWRSLSNNLVIDQNGNVEMSSQNI 547  
 OY 506 DTD-VLBAEADGMVNDALYANKASJVAIFIIDANSGRBLFESTAIDLWAMPG 564  
 DB 548 RIEKPVDAHS--IYESLYRHKSTLETFRMKONGQSMKFRIDACEVL----G 604  
 OY 565 AVSKAEM---LIEKCRMADMLNGDGKVLDNLNEFLAFRLSDLRKE 604  
 DB 605 HIDSQDNKLANTIMDNGSIDPNBLKAFTV--VRYE 642

KW protein phosphatase EF hands-1; RAB21;  
 KW TNF receptor-associated factor 6-binding protein; spaghetti protein;  
 KW HSKM-B; cainexin-t; nexin 3; glia-derived nexin precursor;  
 KW microtubuli-associated protein 4; CDC10;  
 KW transcription co-activator CRSP150; JERKY; SOUL protein;  
 KW tRNA intron-encoded homing endonuclease;  
 KW epithelial cell growth inhibitor; glutathione transferase A4; CGI-69;  
 KW differential display; chronic pain; gene therapy.  
 OS Homo sapiens.

PN WO2004070333-A2.  
 XX  
 PD 19-AUG-2004.  
 XX  
 PP 02-DEC-2003; 2003WO-EP013597.  
 XX  
 PR 05-FEB-2003; 2003DE-01004680.  
 XX  
 PA (CHERF ) GRUENENTHAL GMBH.  
 XX  
 PT Weihe E, Bieller A, Schaefer MK;  
 XX  
 WPI; 2004-604492/58.  
 DR N-PSDB; ADT66672.  
 XX  
 PT New polynucleotides and polypeptides involved in pain regulation, useful  
 PT in screening for pain-regulating compounds for therapy and diagnosis.  
 PS  
 XX  
 CC This invention describes novel polynucleotides and polypeptides involved  
 CC in pain regulation which are useful in screening for pain-regulating  
 CC compounds for therapy and diagnosis. Identifying pain-regulating  
 CC substances comprises incubating a test compound with a cell, and/or  
 CC preparation of a cell, that synthesises at least one specific peptide or  
 CC protein, then measuring either binding of test compound to the protein or  
 CC a functional parameter that is altered by the binding. The protein is one  
 CC or more of any of lutzp, PGP-19, phosphatidylinositol synthase, valosin-  
 CC containing protein, interleukin 6 receptor subunit beta, aspartate  
 CC aminotransferase, neuronal immediate early gene, heat shock protein 27,  
 CC HSC70, calmodulin, syntaxin binding protein 1, splicosomal protein Sip  
 CC 155, neurodop 1, lamcan, leukotriene A4 hydrolase, chondromodulin 1  
 CC Msb4/phosphatidylinositol-4-phosphate 5 kinase, 26S proteasome Ub p12 or  
 CC Ub 2 or UE p44, ingainsin, 3-hydroxy-3-methylglutaryl-Co A synthase,  
 CC phosphoglycerate kinase, RNA polymerase II TP SII-p18 US, protein  
 CC phosphatase BP hands 1, RAB21, TNF receptor-associated factor 6-binding  
 CC protein, ortholog of the Drosophila gene, spaghettii, protein with 4-  
 CC transmembrane domains member 3, vacuolar protein sorting protein, HSKM-B,  
 CC calnexin-t, sorting nexin 3, glia-derived nexin precursor, microtubuli-  
 CC associated protein 4, CDC 10, transciption co-activator CRSP150, JERKY,  
 CC KM-209528, TP SOX10, SOTL protein, tRNA intron-encoded homing  
 CC endonuclease, epithelial cell growth inhibitor, glutathione transferase  
 CC A4 and/or CCI-69, the specified nucleic acids and proteins have been  
 CC identified using the restriction fragment differential display PCR method  
 CC (Biochem. Biophys. Res. Commun., 234 (1997) 16), applied to rats in which  
 CC arthritis has been induced by injection of Complete Freund's adjuvant.  
 CC Compounds identified by the new method, also specified polynucleotides,  
 CC their antisense sequences, encoded proteins and peptides, vectors,  
 CC antibodies and cells, are useful as therapeutic and diagnostic agents,  
 CC particularly for treatment of chronic pain, including by in vivo or in  
 CC vitro gene therapy. This sequence represents the human protein  
 CC phosphatase, EF hands-1 protein.  
 XX  
 SQ sequence 653 AA;

Query Match 36.9%; Score 1271.5; DB 8; Length 653;  
 Best Local Similarity 40.3%; Pred. No. 1.2e-110;  
 Matches 258; Conservative 129; Mismatches 204; Indels 49; Gaps 11;

Oy 3 ENRAAATPIKQYKMRHARREMRORCWCWQIQLNEYASEQPQAELEYKPFNDLIKHMQA 62  
 14 :|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:  
 XX  
 PN WO2003040325-A2.  
 XX  
 PD 15-MAY-2003.  
 XX  
 PP 05-NOV-2002; 2002WO-US035464.

QY 63 AGRKNOQGSANVSVLDDKD--DLVE--ERGDIVA-KIELPIRKHHIDLLDVVERKRG 117  
 :|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:  
 Db 74 HKEELSLRNOSLESEQDMRDRWYDVSIDVPDSYNGPRQLPFLCTIDIDLLBKFQQ- 132  
 QY 118 NRLHPKVVALIREAAKSILKOLPNISPVSTAVSQVTCQVCGSLHGKLDDLVTLHKGUPS 177  
 :|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:  
 Db 133 -ILHARYVLEVLFETKKVVKOMPNTHTQPSPKEVITICGDLHGKLDDLFLLFYKNGLPS 191  
 QY 178 SSNPVTTENGFDVDRGKGEVLILLISLILAFPAVFLNRGHEDSWNARYGPREVES 237  
 :|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:  
 Db 192 ERNPVTFENGFDVDRGKTSRBLIMLUCSPLTYPNDAHLNLNGHEPMNNLRGYFTGLH 251  
 :|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:  
 Db 238 KYPNRHKRILAFIDEVRWLPGSVLNRSVLTVGCGFSSTSLSLDLTKSIRDGYVSLRP 297  
 :|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:  
 Db 252 KYKLHSKRILQILBERFAMWLTIGITIVDNEBLVINGGSESTDNLNLRVNRKMSVLP 311  
 QY 298 PL-----TDGP----LDKTEWQOITPDIMNSDPOATMGCV 328  
 :|||:|||:|||:|||:|||:|||:|||:|||:|||:  
 Db 312 PTEINRHDHTDSKHNKVGVTNAHGRKIKTNGSPTEHLSHEWEQOIDLWSDPRKGNCF 371  
 QY 329 PNTLIGAGWGGPDYTDNFPLCHRYSVUTSHECKNGHFMHDNKLITFSASNYAIG 388  
 :|||:|||:|||:|||:|||:|||:|||:|||:  
 Db 372 PNTCRSGCYQDPDVTSKILNQYQQLNQHHECKEGYECHDGKVUTFSASNYAEG 431  
 QY 389 SNKGAVTRINNQOLMPHFWQVISAASOTKRLSP---KORMGIVESSAKELAVMRDHRDE 445  
 :|||:|||:|||:|||:|||:|||:|||:|||:  
 Db 432 SNRGAVYKLGCGSTTPRFQY---QVTKATCOPQRGRVMTNSAKLRLRVSIRKSD 487  
 QY 446 LDEBFKRYDKPSQGYTSISHCKVMENTVKGPLWHLRDKLAPGFDQSOKVNRTIDL 505  
 :|||:|||:|||:|||:|||:|||:|||:  
 Db 488 UTRAFQDHRKSGKLSVSOWAPCMENILGQNLPMWSLSSNLVNIDONGNVEMSSQNI 547  
 ID ADJ95096 standard; protein, 613 AA.  
 XX  
 AC ADJ95096;  
 XX  
 DT 06-MAY-2004 (first entry)  
 DB Novel NOVX protein sequence #162.

OS antidiabetic; anorectic; cardiotonic; hypotensive; antiarteriosclerotic;  
 KW anorectic; viricide; antibacterial; fungicide; protozoocide; nectropic;  
 KW neuroprotective; anti-parkinsonian; anticonvulsant; osteopathic;  
 KW antiarthritic; anti-inflammatory; dermatological; antiasthmatic;  
 KW antilipemic; gene therapy; metabolic disorder; diabetes; obesity;  
 KW infectious disease; anoxia; cancer; cardiovascular disease;  
 KW hypertension; atherosclerosis; neurodegenerative disorder;  
 KW Alzheimer's disease; Parkinson's disease; epilepsy; immune disorder;  
 KW osteoarthritis; hematopoietic disorder; inflammatory skin disorder;  
 KW asthma; dyslipidemia; neurogenesis; cell differentiation; cell  
 KW proliferation; hematopoiesis; wound healing; angiogenesis;  
 KW chromosome mapping; tissue typing; pharmacogenomic.

OS Homo sapiens.

XX  
 PN WO2003040325-A2.  
 XX  
 PD 15-MAY-2003.  
 XX  
 PP 05-NOV-2002; 2002WO-US035464.

The invention relates to novel isolated polypeptides, mature forms of these, or a sequence that is at least 95 % identical to, or having one or more conservative amino acid substitutions in the polypeptides. The polypeptides, nucleic acid molecules and antibodies are useful in the manufacture of a medicament for treating a syndrome associated with a human disease, preferably a NOX-associated disorder. The nucleic acid molecules, polypeptides and antibodies are useful for treating, preventing or diagnosing diseases such as metabolic disorders, diabetes, obesity, infectious diseases (viral, bacterial, fungal, helminthic, and protozoal), anorexia, cancer, cardiovascular diseases (hypertension, atherosclerosis), neurodegenerative disorders, Alzheimer's disease, Parkinson's disease, epilepsy, immune disorders (osteoarthritis), hematopoietic disorders, inflammatory skin disorders, asthma, and various dyslipidemias. The nucleic acids and polypeptides may also be used as targets for the identification of small molecules that modulate or inhibit e.g. neurogenesis, cell differentiation, cell proliferation, hematopoiesis, wound healing and angiogenesis, in gene therapy, in generation of antibodies that bind immunospecifically to NOX substances for use in therapeutic or diagnostic methods. The nucleic acids are further used as hybridization probes, in chromosome mapping, tissue typing, preventive medicine, and pharmacogenomics. This sequence

Typing, prevention

Query Match 33.8%; Score 1164.5; DB 7; Length 613;  
 Best Local Similarity 39.3%; Pred. No. 1.5e-100;  
 Matches 247; Conservative 115; Mismatches 200; Indels 67; Gaps 13;

05-NOV-2001;	2001US-0338636P.
06-NOV-2001;	2001US-0333072P.
09-NOV-2001;	2001US-0342610P.
15-NOV-2001;	2001US-0335610P.
16-NOV-2001;	2001US-0338532P.
20-NOV-2001;	2001US-0331630P.
20-NOV-2001;	2001US-0331611P.
21-NOV-2001;	2001US-0332152P.
27-NOV-2001;	2001US-0334562P.
28-NOV-2001;	2001US-0333912P.
28-NOV-2001;	2001US-0334020P.
29-NOV-2001;	2001US-0334300P.
30-NOV-2001;	2001US-0334421P.
30-NOV-2001;	2001US-0334562P.
04-DEC-2001;	2001US-0334572P.
04-DEC-2001;	2001US-0336644P.
07-DEC-2001;	2001US-0338314P.
07-DEC-2001;	2001US-0338330P.
10-DEC-2001;	2001US-0339006P.
10-DEC-2001;	2001US-0339008P.
11-DEC-2001;	2001US-0339266P.
01-FEB-2002;	2002US-0353280P.
01-FEB-2002;	2002US-0353288P.
04-FEB-2002;	2002US-0354322P.
04-FEB-2002;	2002US-0354339P.
04-FEB-2002;	2002US-0354403P.
27-FEB-2002;	2002US-0359944P.
05-MAR-2002;	2002US-0361172P.
05-MAR-2002;	2002US-0361183P.
05-MAR-2002;	2002US-0361192P.
05-MAR-2002;	2002US-0362230P.
13-MAR-2002;	2002US-0362652P.
13-MAR-2002;	2002US-0364000P.
13-MAR-2002;	2002US-0364131P.
13-MAR-2002;	2002US-0364132P.
13-MAR-2002;	2002US-0364177P.
17-MAY-2002;	2002US-0381621P.
28-MAR-2002;	2002US-0383675P.
17-JUL-2002;	2002US-0396703P.
06-AUG-2002;	2002US-0401522P.
07-AUG-2002;	2002US-0401594P.
07-AUG-2002;	2002US-0401787P.
15-AUG-2002;	2002US-0403619P.
20-AUG-2002;	2002US-0404821P.
23-AUG-2002;	2002US-0405368P.
23-AUG-2002;	2002US-0405426P.
23-AUG-2002;	2002US-0405631P.
26-AUG-2002;	2002US-0406125P.
04-NOV-2002;	2002US-00287226.

New isolated NOVX polypeptides and polymucrotides, useful for preventing, diagnosing or treating NOVX-associated disorders, e.g. osteoarthritis, obesity, atherosclerosis, cancer, Parkinson's disease, Db 519 -TIVSTLTRYRSDPDEIFNAIDPHSGLISVERFRAMWKLPFSSHYNVIDDSQWNKLA 575



QY	178 SSNIYVFNGLPVDGKRGKGLFLVLLLSSLYAFLPNAFLNRRGHEDSVMMARYSPIREVES	XX
Db	192 BRNFYFNGDFDVKGSNSTEILMLCVSFLVYVNDLHNGRNHDPMMLRYGFTKEIH	PT
QY	238 KYPRHKRKLAFIPIBVRVYRMP-----LASVLSNLVRLVHGGFSDSTSIDL	PT
Db	252 KYKLGKGRQLIQLRFYAMLPTEETNRDHGTDKSKNKVGVTFNA-----HGRIK-----	PT
QY	283 IKSIDRGKVKVSLRPLPQGP--LDKTHWQOIFDMDPQATMGCVNTLRGAGWP	299
Db	300 -----TNGSPTEHLTEHWBQDIDWSPDGKNGEPNTTERGGCVP	342
QY	340 GPDVTDNFLQPHRSYVIRSHHECKENGHRPHMDNKLITISASNYTAIGNSKGAYIRLN	399
Db	343 GPDVTSKLKXQXQMLKSHCPEGBVICHDSKVTTIASNYTGGNSNRGAYIKLCS	402
QY	400 QLMFHPOVISAQSOTKRISF--KORMGIVVESSALKELAVRMNDHBRBLEDSPRKVPK	456
Db	403 GTTPRPFQI---QVTKATCFQPLQRVDMTENSAKILRKERVISRKSDITRAFOLOHR	458
QY	457 DSGYISISHWKMVENUTVGLPWLRLXKAPOTSQKMYRTLDLTD-VILERA	515
Db	459 KSKGLSVSOMAFCHENILGILNPWRSLSSTNLVNDONGWYMSFONTRIEKPVORAH	518
QY	516 DGMSWDAKYANKASLVAFFNIIDADNSCETTUDBEFTADLLVLAHMPGAYSKAEMLERKC	575
Db	519 --TLVETLWYRSDLBIFNAIDTDHSGLISUEFRAMWKLFSSHYNHVIDDSQVNKL	575
QY	576 RMDMGLDCKVQDNEBLFARFLSDLRHK	604
Db	576 NIMDLNKDGSSIDFNEBLKAFYV--VHRYE	602
<b>RESULT 7</b>		
ID	ADY05391	standard; protein; 488 AA.
AC	ADY05391	
XX	21-APR-2005 (first entry)	
DE	Plant full length insert polypeptide seqid 61206.	
KW	plant protectant; Plant growth regulator; gene therapy; plant; recombinant DNA construct; physical array; plant breeding marker; cold tolerance; heat tolerance; drought tolerance; herbicide tolerance; extreme osmotic condition; pathogen tolerance; pest tolerance; growth rate; cell cycle pathway; disease resistance; plant; yield; plant growth; plant development; seed oil; protein yield; protein content.	
OS	Unidentified.	
PN	US2004034888-A1.	
XX	19-FEB-2004.	
PP	28-APR-2003; 2003US-00425114.	
PD		
XX		
<b>RESULT 8</b>		
PR	06-MAY-1999; 99US-00304517.	ABB63840
PR	05-NOV-2001; 2001US-00985678.	ABB63840 standard; protein; 520 AA.
XX		
PA	(LIU J.) LIU J.	AC
PA	(ZHOU/) ZHOU Y.	ABB63840;
PA	(KOVALIC) KOVALIC D. K.	XX
PA	(SCREB/) SCREB S. E.	DT
PA	(TABA/) TABASKA J. B.	26-MAR-2002 (first entry)
XX	(CAO Y.) CAO Y.	DB
PT	Liu J., Zhou Y., Kovalic DK, Screen SE, Tabaska JB, Cao Y;	XX
XX		KW
DR		pharmaceutical.
WPI; 2004-180133/17.	OS	Drosophila melanogaster.
XX		

XX	W0200171042-A2.	ABG09989
PD	27-SEP-2001.	standard; protein; 500 AA.
XX		ID ABG09989
PP		AC ABG09989;
XX		XX ABG09989,
PR	23-MAR-2001; 2001WO-US009231.	AC ABG09989;
XX		XX ABG09989,
PR	23-MAR-2000; 2000US-0191637P.	DT 13-FEB-2002 (first entry)
PR	11-JUL-2000; 2000US-00614150.	DE Novel human diagnostic protein #9980.
XX		XX Human; chromosome mapping; gene mapping; gene therapy; forensic; food supplement; medical imaging; diagnostic; genetic disorder.
PA	(PBKR ) PE CORP NY.	KW
XX	Venter JC, Adams M, Li PWD, Myers EW;	KW
XX	PT DR N-PSDB; ABL07943.	XX PI Homo sapiens.
XX	PT WPI; 2001-656860/75.	XX OS
XX	PT N-PSDB; ABL07943.	XX PN WO200175067-A2.
XX	PT New isolated nucleic acid detection reagent for detecting 1000 or more genes from Drosophila and for elucidating cell signaling and cell-cell interactions.	XX PD 11-OCT-2001.
CC	The invention relates to an isolated nucleic acid detection reagent capable of detecting 1000 or more genes from Drosophila. The invention is useful in developmental biology and in elucidating cell signalling and cell-cell interactions in higher eukaryotes for the development of insecticide, therapeutics and pharmaceutical drugs. The invention discloses genomic DNA sequences (ABL1676-ABL3051), expressed DNA sequences (ABL180-ABL16175) and the encoded proteins (ABP5773-ABP72072). The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at <a href="ftp://wipo.int/pub/published_pct_sequences">ftp://wipo.int/pub/published_pct_sequences</a>	XX PP 30-MAR-2001; 2001WO-US008631.
CC	XX	XX PR 31-MAR-2000; 2000US-00540217.
CC	XX	XX PR 23-AUG-2000; 2000US-00649167.
CC	XX	XX PA (HYSB-) HYSEQ INC.
CC	XX	XX PT Drmanac RT, Liu C, Tang YT;
CC	XX	XX DR WPI; 2001-639362/73.
CC	XX	XX N-PSDB; AAS74176.
CC	CC New isolated polynucleotide and encoded polypeptides, useful in diagnostics, forensics, gene mapping, identification of mutations responsible for genetic disorders or other traits and to assess biodiversity.	XX PT
CC	CC	XX PT
SQ	Sequence 520 AA;	XX PT
Query Match	16.8%; Score 577; DB 4; Length 520;	XX PS Claim 20; SEQ ID NO 40348; 103pp; English.
Best Local Similarity	32.1%; Pred. No. 5.5e-45;	XX XX
Matches	136; Conservative 83; Mismatches 153; Indels 52; Gaps 15;	CC The invention relates to isolated polynucleotide (I) and polypeptide (II) sequences. (I) is useful as hybridisation probes, polymerase chain reaction (PCR) primers, oligomers, and for chromosome and gene mapping and in recombinant production of (II). The polynucleotides are also used in diagnostics as expressed sequence tags for identifying expressed genes. (I) is useful in gene therapy techniques to restore normal activity of (II) or to treat disease states involving (II). (II) is useful for generating antibodies against it, detecting or quantitating a polypeptide in tissue, as molecular weight markers and as a food supplement. (II) and its binding partners are useful in medical imaging of sites expressing (II). (I) and (II) are useful for treating disorders involving aberrant protein expression or biological activity. The polypeptide and polynucleotide sequences have applications in diagnostics, forensics, gene mapping, identification of mutations responsible for genetic disorders or other traits to assess biodiversity and to produce other types of data and products dependent on DNA and amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic amino acid sequences of the invention. Note: The sequence data for this patent did not appear in the printed specification, but was obtained in electronic format directly from WIPO at <a href="ftp://wipo.int/pub/published_pct_sequences">ftp://wipo.int/pub/published_pct_sequences</a>
Qy	4 NAIRAA-TIQKWMRHRQARREM--QRCNWQIQFQNLYEYASQQDQEYKPFNDLK-	CC CC
Db	108 SAVKDPATLGKYVRAAAMSISQFKQALCDERFAKCR -PMDKAULKFTCNKVYM 166	CC CC
Qy	58 -----HMPQAAGRKNQYQGSYHVSIDD-KDLDVERFGDTWAKIELPIRKHNIDL 107	CC CC
Db	167 RAFAIAVADKPEKT-LSEMTSDMENITBDDYKQPLSED----GKTLKFMKE---- 215	CC CC
Qy	108 LIDYRKKGNRNLRPKYVALLTREAKSLKOLPNISPVNSTAVSQCQTYCGDLHKLDDLL 167	CC CC
Db	216 LMERYKAOK--RLRKRFAYKILCBLDTYMAQPSLVDITWPDEEKKPCTGDDHQFYDLM 273	CC CC
Qy	168 WLHKHGLJSNSNTYTFVNGDFDQKRGKGLLVLILLSLYLAFFPWAFLNQGHLRHSVNA 227	CC CC
Db	274 NIFELINGLSEKNPFLNFQDFDGGSFSTCIPFQKLLYPHPPFLARGNHSINNOQ 333	CC CC
Qy	228 RYGFTREVSKYPRNHRKTLAFIDEVRYWRMLPGSVLNSRVLIVHGGR--SDSTSIDLTKS 285	CC CC
Db	334 MYGTTGEVTAKYT--SAMADIFTQYFVNWLPLCHICINOKLVMMEGGLFSTDYVLDHRR 390	CC CC
Qy	286 IDRQKXVSLRPLPDLGEPUDKTEMQFIDIMMSDPOATMGCVPTNLRGAGWFGPDDYD 345	CC CC
Db	391 IERN----CQPP-----BEGMCBLWSPQOMG-LGCKSKRKGVIQGPDVTE 434	CC CC
Qy	346 NFLQBRHLSVIRSHCCKNGHBFHMHDNKITITISASNTAIGNSNGAYRL-ANQLMFH 404	CC CC
Db	435 KFCKDUNLQDYLIRSHIEVKONGYEVHNGKTCITVFSAPVYCDTMENGMAPITCNIKEN 494	CC CC
Qy	405 FVQY 408	CC CC
Db	495 YKSF 498	CC CC
SQ	Sequence 500 AA;	XX SQ
Query Match	16.6%; Score 571; DB 4; Length 500;	XX
Best Local Similarity	34.8%; Pred. No. 1.9e-44;	XX
Matches	143; Conservative 73; Mismatches 153; Indels 42; Gaps 16;	Qy 11 PIQKWMRHRQARREM-QRCNWQIQFQNLYEYAS-EQDQEYKPFNDLKHMPOAAGRN 67
Db	99 YIMGYTTRASNNMALKFRAALRDYETVVKPKHDKAOKYKOCNKNVKT--PGRPER-- 154	Db 99 YIMGYTTRASNNMALKFRAALRDYETVVKPKHDKAOKYKOCNKNVKT--PGRPER-- 154
Qy	68 QYQCSAHV-SVLDKDD---DLVERFG--DIVAKIELPIRKHNIDLLDVFRKGNRL 120	Qy



PA (RIGE-) RIGEL PHARM INC.

XX

PT Ku X, Luo Y;

XX

DR WPI; 2001-55775/62;

XX

N-PSDB; AAD16789, AAD16790.

XX

Cell cycle protein having phosphatase activity for generating agonist and antagonist e.g. antibodies of the polypeptides, both of which are useful in gene therapy and in screen assays.

XX

PS Disclosure; Fig 2, 75pp; English.

XX

The patent discloses cell cycle proteins and nucleic acids encoding them. The cell cycle proteins are capable of promoting cell cycle progression or cell cycle arrest. They are capable of promoting or inhibiting DNA repair. The invention also provides methods for screening a bioactive agent capable of modulating the cell cycle or interfering with the binding of a cell cycle protein, protein phosphatase type 5 (PP5). The method involves combining a cell cycle protein and a candidate bioactive agent and determining the effect on the cell in the presence or absence of the candidate agent. Cell cycle proteins bind to Rad9 proteins. They are useful for generating an agonist and antagonist of the cell cycle proteins, both of which are useful in gene therapy and in screen assays. They are useful in the diagnosis and treatment of cancer. The proteins of the invention are used as vaccines. They are useful for generating polyclonal or monoclonal antibodies to the proteins. The present sequence is a novel cell cycle protein, protein phosphatase type 5 (PP5) of the invention.

XX

SQ Sequence 499 AA;

Query Match 16.6%; Score 570; DB 4; Length 499;

Best Local Similarity 33.7%; Pred. No. 2.4e-44; Matches 143; Conservative 71; Mismatches 142; Indels 68; Gaps 16;

CC 11 FTKWYRHQRREMRQRCWQIQFQMLKYSRQDQELEYKFENDLKHMPAAGRKQYQ 70

CC 95 YIKGYYRRRAASNMGK-----FR-----AALRD-----YETVVKVKPDKDAKMYQ 137

CC 71 -----GSAHV-SVLDDKD---DLVERFG--DIVNAKIEPIRKONHDL 107

CC 138 ECKMKIVKOKAPERAIAQDEHKSKSVWDSDIESTNTIRBYSGPKLEDEKUTSFMKR--- 193

CC 108 LIDVFKRKGKGNRLHPKYVALLRERAKSLSKQJQNISPVSTAVSQVTWCGLAIGKDDLL 167

CC 194 IMKWYKDKK---KLHRKCAQYQLVQKVKEVLKSLTVEETLKSTEKITVCGDTHQFDLL 251

CC 168 VVLAHKGSPSSPYVFGDPYDVGKQGLEVUILLISLYAFPNAVLNRAHRHDSTMNA 227

CC 252 NTFELNGLIPSETNPYIENGDFVRGFSFSEVILTLFGPKLYPDHFHLRKGHETNMQ 311

CC 228 RICFIREVESKVRPHKELAFDEVRMPLGVSUNSRVLUVHGGP--SDSSTDLIKS 285

CC 312 IYGRBGEVTAKYT---ADMVELPSEVFPLAQCINGKVLIMHGULFSBBDVTDIHK 368

CC 286 IDRGKYSVILRPLTDGEPPLDKTEWQOQFDIMMSDPOATMGCVPNTLRAGA3WMFGPDVTD 345

CC 369 IERNR---QPP---DSCPM----CDLIMSDPQPGNG-RSISKGVSCQFGPDVTK 412

CC 346 NFYQHRHSYVRSHECKPNGFHMNDKLTIFSAINTYAGGSNKAYIRU-NNGLMPH 404  
413 AFLBENNLDYIIRSHEVKGAEVAGSGRCVTVPSAPNCYDONGNKASYIHLQSDLRPQ 472

CC 405 FVQY 408  
DB 473 FHQF 476

AC ADB64103;

XX

DT 29-JAN-2004 (first entry)

XX

DS Human Protein P53041, SEQ ID NO 10051.

XX

KW Human; pain; neuronal tissue; gene therapy'; spinal segmental nerve injury; chronic constriction injury; CCI; spared nerve injury; SNI; Chung.

XX

OS Homo sapiens.

XX

PN WO2003016475-A2.

XX

PD 27-FEB-2003.

XX

PP 14-AUG-2002; 2002WO-US025765.

XX

PR 14-AUG-2001; 2001US-0312147P.

PR 01-NOV-2001; 2001US-0346382P.

PR 26-NOV-2001; 2001US-0333347P.

XX

PA (GBHO ) GEN HOSPITAL CORP.

XX

PA (FARB ) BAYER AG.

XX

PI Woolf, C, D'urso, D, Befort, K, Costigan, M;

XX

WPI; 2003-268312/26.

XX

DR GENBANK; P53041.

XX

PT New composition comprising two or more isolated polypeptides, useful for preparing a medicament for treating pain in an animal.

XX

PS Claim 1; Page; 1017pp; English.

XX

CC The invention discloses a composition comprising two or more isolated rat or human polynucleotides or a polynucleotide which represents a fragment, derivative or allelic variation of the nucleic acid sequence. Also claimed are a vector comprising the novel Polynucleotide, a host cell comprising the vector, a method for identifying a nucleotide sequence which is differentially regulated in an animal subjected to pain and a kit to perform the method, an array, a method for identifying an agent that increases or decreases the expression of the polynucleotide sequence that is differentially expressed in neuronal tissue of a first animal subjected to pain, a method for identifying a compound which regulates the expression of a polynucleotide sequence which is differentially expressed in an animal subjected to pain, a method for identifying a compound that regulates the activity of one or more of the polynucleotides, a method for producing a pharmaceutical composition, a method for identifying a compound or small molecule that regulates the activity in an animal of one or more of the polypeptides given in the specification, a method for identifying a compound useful in treating pain and a pharmaceutical composition comprising the one or more polypeptides or their antibodies. The polynucleotide or the compound that modulates its activity is useful for preparing a medicament for treating pain (e.g. spinal segmental nerve injury (Chung), chronic constriction injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene therapy). The sequence presented is a human protein (shown in Table 2 of the specification) which is differentially expressed during pain. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic form directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences.

XX

SQ Sequence 499 AA;

Query Match 16.6%; Score 570; DB 7; length 499;

Best Local Similarity 33.7%; Pred. No. 2.4e-44; Matches 143; Conservative 71; Mismatches 142; Indels 68; Gaps 16;

CC 11 FTKWYRHQRREMRQRCWQIQFQMLKYSRQDQELEYKFENDLKHMPAAGRKQYQ 70

CC 95 YIKGYYRRRAASNMGK-----FR-----AALRD-----YETVVKVKPDKDAKMYQ 137

QY	71 -----GSANV-SVLDKD---DLVERFG--DIVAKIELPLTRKNHDL	107	CC	overexpressed in cancer tissues compared to normal tissues, and may thus serve as effective targets for the diagnosis and treatment of cancer in mammals. The invention also relates to nucleic acid and polypeptide sequences at least 80% identical to the TAT nucleic acids and polypeptides; expression vectors and host cells comprising a TAT nucleic acid; an antibody specific for a TAT polypeptide; a peptide or organic molecule which binds to a TAT polypeptide; fusion proteins comprising a TAT polypeptide; and methods and compositions for the treatment or diagnosis of cancer in mammals. TAT polypeptides, nucleic acids and antibodies, antagonists, binding molecules and compositions are useful for diagnosing or treating a cell proliferative disorder associated with increased TAT expression, particularly cancers such as breast cancer, colorectal cancer, lung cancer, ovarian cancer, liver cancer, bladder cancer, pancreatic cancer, cervical cancer, cancers of the central nervous system, melanoma and leukaemia. TAT nucleic acids may further be used as hybridisation probes, in chromosome and gene mapping, in chromosome identification and in gene therapy. The present sequence represents a TAT polypeptide of the invention
Db	138 ECKNIVKQKAFAERIAQDGDKRVSVDSDIESMTIEVSGPKLBDGKVTTSMKE---	193	CC	
QY	108 LIDVFRKGNGNLAPKVVALLIREAAKSQKLQNLNPISPVSTAVSQVTGCGDQIUGKLDDL	167	CC	
Db	194 LMQWYKDK--KLHKCAQVQLVQEVKLSKLISTIVETLKETKTVCGDQIUGKLDDL	251	CC	
QY	168 WUHKNGLSSSNPVPYENGDFVDRGKRGLEVLLSXLAPPNAFLVNGGF--SDSTSLLDKS	285	CC	
Db	312 IYGFEGEVKAKT--AQMYELSERFENPLAQCINGKLUMIGGLPSBDGVLDDIRK	368	CC	
QY	286 IDRKVVSILRPLTDGEPBLDKTEWQOQFDIMWDSDPDTMGCVPNTLGGAGVWFGPVDYD	345	CC	
Db	369 IERUR----OPP--DSGPW----CDLMSDPQFONG-RSISKRCVSCQGPDVTK	412	CC	
QY	346 NFIQRHRISYVRSHECKPGHFMHMKITIFSAINTYAGSNKAYTRL-NNQLMHF	404	CC	
Db	413 AFLENNLDYIIRSHEVKAGEYVAH3GRCVTVFSAPNYCDOMGKASYIHLQGSDLRQ	472	XX	
RESULT 13				
ABM82127			Query Match	16.6%; Score 570; DB 8; Length 499;
ID	ABM82127 standard; protein: 499 AA.		Best Local Similarity	33.7%; Pred. No. 2 4e-44; Mismatches 143; Conservative 71; Mismatches 142; Indels 68; Gaps 16;
XX			Matches	143; Conservatve 71; Mismatches 142; Indels 68; Gaps 16;
AC			Db	11 FICKWYRHQARREMRORRNQWQFQNLWYASEQDQALYKFENDLKMPOAGKRYQ 70
XX			QY	95 YIKGYRRRAASNMALGK-----FR-----AALRD-----YETVVKTPHDOKAOXQ 137
DT	18-NOV-2004 (first entry)		Db	71 -----GSANV-SVLDKD---DLVERFG--DIVAKIELPLTRKNHDL
XX			QY	138 ECKNIVKQKAFAERIAQDGDKRVSVDSDIESMTIEVSGPKLBDGKVTTSMKE---
DE	Tumour-associated antigenic target (TAT) polypeptide PRO25881, SEQ:5498.		Db	194 LMQWYKDK--KLHKCAQVQLVQEVKLSKLISTIVETLKETKTVCGDQIUGKLDDL
XX			QY	168 WUHKNGLSSSNPVPYENGDFVDRGKRGLEVLLSXLAPPNAFLVNGGF--SDSTSLLDKS
KW	Tumour-associated antigenic target; TAT; human; overexpression; cancer;		Db	312 IYGFEGEVKAKT--AQMYELSERFENPLAQCINGKLUMIGGLPSBDGVLDDIRK
KW	tumour; diagnosis; cell proliferative disorder; breast cancer;		QY	286 IDRKVVSILRPLTDGEPBLDKTEWQOQFDIMWDSDPDTMGCVPNTLGGAGVWFGPVDYD
KW	colorectal cancer; lung cancer; ovarian cancer; liver cancer;		Db	369 IERUR----OPP--DSGPW----CDLMSDPQFONG-RSISKRCVSCQGPDVTK
KW	central nervous system cancer; bladder cancer; pancreatic cancer;		QY	346 NFIQRHRISYVRSHECKPGHFMHMKITIFSAINTYAGSNKAYTRL-NNQLMHF
KW	cervical cancer; melanoma; leukaemia; hybridisation probe;		Db	413 AFLENNLDYIIRSHEVKAGEYVAH3GRCVTVFSAPNYCDOMGKASYIHLQGSDLRQ
KW	chromosome identification; chromosome mapping; gene mapping;		QY	404
KW	gene therapy; cytostatic.		Db	472
OS	Homo sapiens.		QY	
XX			Db	
PN	WO2004030615-A2.		QY	
XX			Db	
PD	15-APR-2004.		QY	
XX			Db	
PR	29-SEP-2003; 2003W0-US028547.		QY	
XX			Db	
PA	(GETH ) GENENTECH INC.		QY	
XX			Db	
PT	Wu TD, Zhang Z, Zhou Y;		QY	
XX			Db	
DR	WPI; 2004-347921/32.		QY	
DR	N-ISDB; ACM40614.		Db	
XX			QY	
PT	New tumor-associated antigenic target polypeptides and nucleic acids, useful in preparing a medicament for treating or detecting a proliferative disorder, e.g. breast, lung, colorectal, ovarian or prostate cancer or tumor.		Db	
PT	Claim 12; SEQ ID NO 5498; 7273PP; English.		QY	
XX			Db	
CC	The invention relates to human tumour-associated antigenic target (TAT) polypeptides, and their related nucleic acids. The TAT polypeptides are		QY	
CC	Rattus norvegicus.		Db	







QY 286 IDRGKYSVILRPLDGPRLDKTEWQQIFDDIMSDPQATMGCVPTNLARGAGWFGPDVTD 345  
 XX ::|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:  
 Db 369 IERNR----QPP--DSGPM-----CDLMSDPQFQNG-RSISRGVSCQFGPDVTK 412  
 QY 346 NFQHRLSYVIRSHECKENGHFRPMHDNKITIFSASNYTGSNKGAYIRL-NNOLMPH 404  
 XX ::|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:  
 Db 413 AFLERENNLDIYIIRSHEVKAEGEVAHGRCVTVFSAPNYCQDMQNKASYIHLQGSDLRQ 472  
 QY 405 FVQY 408  
 XX ::|:|:  
 Db 473 FHQP 476

RESULT 19

AAB09732 standard; protein; 499 AA.

XX AAB09732;

DT 29-NOV-2001 (first entry)

DE Protein phosphatase type 5 (PP5) variant, D274A/R275A.

XX Cell cycle protein; DNA repair; protein phosphatase type 5; PP5; gene therapy; cancer; vaccine; cytostatic; vulnerable; antiinflammatory; mutant; mutein; variant.

OS Unidentified.

FH Key Location/Qualifiers

FT MISC-difference 274 /note= "Wild-type Asp is substituted with Ala"

FT MISC-difference 275 /note= "Wild-type Arg is substituted with Ala"

PN WO200164913-A2.

XX PR 02-MAR-2001; 2001WO-US006849.

XX PR 02-MAR-2001; 2000US-00517779.

XX PA (RIGE-) RIGEL PHARM INC.

XX PI Xu, Luo Y;

XX DR WPI; 2001-557775/62.

XX PT Call cycle protein having phosphatase activity for generating agonist and antagonist e.g. antibodies of the polypeptides, both of which are useful in gene therapy and in screen assays.

PS Claim 16, Page: 75pp; English.

The patent discloses cell cycle proteins and nucleic acids encoding them. The cell cycle proteins are capable of promoting cell cycle progression or cell cycle arrest. They are capable of promoting or inhibiting DNA repair. The invention also provides methods for screening a bioactive agent capable of modulating the cell cycle or interfering with the binding of a cell cycle protein, protein phosphatase type 5 (PP5). The method involves combining a cell cycle protein and a candidate bioactive agent and determining the effect on the cell in the presence or absence of the candidate agent. Cell cycle proteins bind to Rad51 proteins. They are useful for generating an agonist and antagonist of the cell cycle proteins, both of which are useful in gene therapy and in screen assays. They are useful in the diagnosis and treatment of cancer. The proteins of the invention are used as vaccines. They are useful for generating polyclonal or monoclonal antibodies to the proteins. The present sequence is D274A/R275A variant of a novel cell cycle protein, protein phosphatase type 5 (PP5) of the invention. Note: The present sequence is not shown in the specification, but is derived from the PP5 protein shown in Fig 2 (AAB09722).

XX Sequence 499 AA:

Query Match 16.2%; Score 556; DB 4; Length 499;

Best Local Similarity 33.3%; Pred. No. 5e-43;

Matches 141; Conservative 71; Mismatches 144; Indels 68; Gaps 16;

Db 11 PIQKWRHQARREMQRNCNWQFQNRYASQDQALYKPFNDLTHMPQAGRKQYQ 70  
 XX 95 VIKGYYRRASMAIGK-----FR-----AALRD-----YETVVKVPHDKAKQYQ 137

QY 71 .....-SSAHV-SVLKD---DLVERFG-DIVVAKIELDPIRKVHDL 107  
 Db 138 BCNKKIVKQKAPERAIADEHKSVVWSVSLDISWTIBYSGPKLDRSKVITSMKR--- 193  
 QY 108 LIDVFRKRKGRRNRLHPKVALTREEAKSLKOLNPNTSPVSTAVSQVTUGDPLRKGKUDLL 167  
 Db 194 LMQWYKOK- KLARKKAVQYLWQVKELISKISTVTTKTEKITVCGGRTGQFVDL 251  
 Db 168 VVLAHKQNSPSSSNPYVNGDFDYGKGLLEVLLLISLYLAPPNAVFLANGMHEDSTNNAA 227  
 Db 252 NIFELNLSPSETNPYVNGDFVAGSISPSVEVILTFEGPKLKPDPDHFLRGRHBTDMNQ 311  
 QY 228 RQFPRIRSKSPKRNHRKILAFIDEVVRMLPLGSVNSRVLVHGGZ--SDTSLDIKS 285  
 Db 312 RYFGEGETKAKTYT---RQMYELFSEVFLWPLAQCINGKVLINHGGFBSBDCVTLDDIRK 368  
 QY 286 IDRGKYSVILRPLDGPRLDKTEWQQIFDDIMSDPQATMGCVPTNLARGAGWFGPDVTD 345  
 Db 349 IERNR----QPP--DSGPM-----CDLMSDPQFQNG-RSISRGVSCQFGPDVTK 412  
 QY 346 NFQHRLSYVIRSHECKENGHFRPMHDNKITIFSASNYTGSNKGAYIRL-NNOLMPH 404  
 Db 413 AFLERENNLDIYIIRSHEVKAEGEVAHGRCVTVFSAPNYCQDMQNKASYIHLQGSDLRQ 472

RESULT 20

ADS44165 standard; protein; 473 AA.

XX ADS44165;

XX DT 02-DRC-2004 (first entry)

DB Bacterial polypeptide #22595.

XX Recombinant DNA construct; transformed plant; improved plant property; cold tolerance; heat tolerance; drought tolerance; herbicide; osmotic; KW pathogen tolerance; pest tolerance; plant disease resistance; KW cell cycle pathway modification; plant growth regulator; seed oil yield; protein yield; carbohydrate; KW homologous recombination; seed oil yield; protein yield; carbohydrate; KW nitrogen; phosphorus; photosynthesis; lignin; galactomannan; KW bacterial polypeptide.

XX OS Bacteria.

XX PN US2003233675-A1.

XX PD 18-DRC-2003.

XX PF 20-FEB-2003; 2003US-00369493.

XX PR 21-FEB-2002; 2002US-0360039P.

XX (CAOY/) CAO Y.

PA (HINK/) HINKLE G J.

PA (SLAT/) SLATER S C.

PA (CHEN/) CHEN X.

PA (GOLD/) GOLDMAN B S.



CC 72 SARVISVLDKID--DLVE-----EGGDIWAKIEFIRKWHIDLI---- 110  
 CC QY DDKAKLNUKLUVECKIVKUQOLAFFAIEKQDLSAAGGIVERSMADSYGVRL 242  
 CC QY 190 -----DNKDKAKLNUKLUVECKIVKUQOLAFFAIEKQDLSAAGGIVERSMADSYGVRL 242  
 CC QY 111 -----VERKKGRNRLAKHRYVALIREAKSLKQLPNTSPVSTAVSOCVTVC 157  
 CC QY 243 EGMENTQEFDMMERBRPKKLKLYVQIIIAWRNTVNEPTVNEVDIPEDVQLTVCG 302  
 CC QY 158 DLKG-----KLDLILUVLHKGLPSSSNPVYENPDYDGRKRGLE 197  
 CC QY 303 DTHGWALAWPPPSKSMWLTLNTCEAQYDFMLERFLRNGFPSPDKHRYLNGFDYDRGSWSTE 362  
 CC QY 198 VLLILSLYLAPNAPVLQRGNHEDSVMDNARYGTREVSKY----- 239  
 CC QY 363 IAIILAYKWLPRPGPFINNGHETDDMRYVGEGECHKYMNRXASRTLCAHYATSKI 422  
 CC QY 240 -----PRWKRLTAFIDEVRYMPPLGSVLSRVLTIVHGCF--SDSPLSLDKS 285  
 CC QY 423 FERTPADTSHLPLRTYK---LFSSEPSALPLATLIGKPLFVLLGGFLSDDNVTDIIRK 478  
 CC QY 286 IDRCKVSLRPPDGTPLDKTEWQKQFDIMSDQATMGCVNTLKGAGVWFGPDYTD 345  
 CC QY 479 LDRHQ---KQPGAG-----LMHEMLWTDQPEPGPGRPSK-KGVGMQGPDPYTK 524  
 CC QY 346 NFLORRHLSYVIRSHEVMDGYEEBHDGKCITVFSAPKCDMTENKGAYINVIGPDYKJPK 584  
 CC QY 525 RFCDDKNGLRAIRSHEVMDGYEEBHDGKCITVFSAPKCDMTENKGAYINVIGPDYKJPK 584  
 CC QY 406 VOY 408  
 CC QY 585 SQP 587  
 RESULT 22  
 ABG09988 standard; protein; 714 AA.  
 AC ABG09988;  
 XX DT 13-FEB-2002 (first entry)  
 DB Novel human diagnostic protein #9979.  
 XX KW Human; chromosome mapping; gene mapping; gene therapy; forensic; food supplement; medical imaging; diagnostic; genetic disorder.  
 XX OS Homo sapiens.  
 XX PN WO200175067-A2.  
 XX 11-OCT-2001.  
 XX PD 30-MAR-2001; 2001WO-US008631.  
 XX PR 31-MAR-2000; 2000US-00540217.  
 XX PR 23-AUG-2000; 2000US-00649167.  
 XX PA (HYSE-) HYSEQ INC.  
 XX PI Drmanac RT, Liu C, Tang VT;  
 XX DR WPI; 2001-639362/73.  
 XX N-FSDB; RAS#4175.  
 PT New isolated polynucleotide and encoded polypeptides, useful in diagnostics, forensics, gene mapping, identification of mutations, PT responsible for genetic disorders or other traits and to assess biodiversity.  
 PT Claim 20; SEQ ID NO 40347; 103pp; English.  
 CC The invention relates to isolated polynucleotide (I) and polypeptide (II) sequences. (I) is useful as hybridisation probes, polymerase chain reaction (PCR) primers, oligomers, and for chromosome and gene mapping, and in recombinant production of (II). The polynucleotides are also used in diagnostics as expressed sequence tags for identifying expressed genes. (I) is useful in gene therapy techniques to restore normal activity of (II) or to treat disease states involving (II). (II) is useful for generating antibodies against it, detecting or quantitating a polypeptide in tissue, as molecular weight markers and as a food supplement. (III) and its binding partners are useful in medical imaging of sites expressing (II). (I) and (II) are useful for treating disorders involving aberrant protein expression or biological activity. The polypeptide and polynucleotide sequences have applications in diagnostics, forensics, gene mapping, identification of mutations and to produce other types of data and products dependent on DNA and amino acid sequences. ABG0010-ABG30377 represent novel human diagnostic sequences of the invention. Note: The sequence data for this patent did not appear in the printed specification, but was obtained in electronic format directly from WIPO at ftp://wipo.int/pub/published\_pct\_sequences  
 CC Sequence 714 AA:  
 Query Match 14.4%; Score 496.5; DB 4; Length 714;  
 Best Local Similarity 29.2%; Prod. No. 3\_9e-37;  
 Matches 140; Conservative 62; Mismatches 147; Indels 131; Gaps 15;  
 QY 41 SEQDQABLYKFNDLLIKMPQAGKRNQYQ-----GSARV-SVLDKDK-  
 DB 231 SEHAAATYLL---WVKOKPHKDAMKQYQECNKIVKQAKPERAIAGDHKRSVVDSDI 287  
 QY 83 --DLVESEFG--DIWNAKIELPIRKHNIDLLDVFRKGRGRNLRHPKVYALLIBAAKSLK 137  
 DB 288 BSMTIEBVSGPKBLDGKVTSFME---LMQWYDOK--KJURKCAVQYLQVKEVL 341  
 QY 138 QLMNPVSTAVSQVTVGDLAKDLYVLUHKNGKLPSSUPY-----  
 DB 342 KLSLTIVETLTKETKIKTVCGDPTGQPYDINTTFLNGPSETMVPVSAAEPPSPHLHQ 401  
 QY 184 -----FNGDPFDGRKGRGLEVLILLISLYAP 209  
 DB 402 PQQPVCTEWTRPPIPGKDTLTERGLAMGHQGOIIFNGPFDVDRSPSPSVEVLTIFGPKLY 461  
 QY 210 PNAVFLNRGNHEDSVMDNARYGTREVSKYPRWKRLTAFIDEVRYMPPLGSVLSRVLI 269  
 DB 462 PDKHFLRGNHETDNMQTYGFGBVKAKYT--AQMVELSFSEFWELPLACINGKVL 518  
 QY 270 VHGCP--SDSTSLLDKSIDRPGKVSILRPPDGTPLDKTEWQKQFDIMSDQ----- 322  
 DB 519 MRCGLFSRDGVTDIICRER----OPP--DSALERRMRGVPLARNNSRSPQPHL 571  
 QY 323 -----ATMCQCVPTNL---RGAQWFGPDYTDNFHQ 349  
 DB 572 LGVFLHLGAVMASTVVTMVRGEGTIVTAHHTCPYLNGRSISKRGVSCQFGPDYTKAFLB 631  
 QY 350 RHLBSYVIRSHRKPKNGHFMNKILITFSNSNYAIGSNKAYIRL-NNQLMRPHQY 448  
 DB 632 ENNDYDITRSHEVKAEGYEVAGGRCVTVFSAPNCIDOMGNKASYIHLQGSLPQPHQ 691  
 RESULT 23  
 ABR53882 standard; protein; 513 AA.  
 ID ABR53882  
 AC ABR53882;  
 XX DT 20-JUN-2003 (first entry)  
 XX DB Protein Sequence #SEQ ID 2629.  
 XX KW Multiprotein complex; eukaryote; drug target; diagnosis.  
 XX OS Saccharomyces cerevisiae.

PN EP1250494-A1.  
 XX  
 PD 20-NOV-2002.  
 XX  
 PR 20-DEC-2001; 2001EP-00130253.  
 XX  
 PR 15-MAY-2001; 2001EP-00111774.  
 XX  
 PA (CELL-) CELIZOME AG.  
 XX  
 PI Bauer A, Gavin A, Grandi P, Krause R, Kruse UD, Kuester BD,  
 XX Marzioch M, Schultz JD, Superti-Furga GD;  
 XX DR WPI; 2003-250078/25.  
 XX N-PSDB; ACC61924.

PT New isolated protein complexes useful for diagnosing a disease or  
 PT preferably, a drug target in the treatment or prevention of disease or  
 PT disorder.  
 XX Disclosure; SEQ ID NO 2629; 17pp + sequence listing; English.  
 CC The invention relates to multiprotein complexes from eukaryotes. Proteins  
 CC of the invention and DNA sequences encoding them are given in records  
 CC ABR5268-ABR5303 and ACC60610-ACC6194 respectively. The complexes are  
 CC obtainable by using a protein as a bait and isolating the set of proteins  
 CC which is attached thereto from cells. Such protein complexes may comprise  
 CC up to 30 distinct proteins. Protein complexes of the invention are useful  
 CC for diagnosing a disease or disorder, or as a target for an active agent  
 CC of a pharmaceutical, preferably a drug target in the treatment or  
 CC prevention of a disease or disorder. Note: The sequence data for this  
 CC patient is not represented in the printed specification, but is based on  
 CC sequence information supplied by the European Patent Office. The complete  
 CC document is available on CD-ROM

XX SQ Sequence 513 AA;

Query Match 13.9%; Score 479.5; DB 6; Length 513;  
 Best Local Similarity 32.3%; Pred. No. 9.3e-36;  
 Matches 130; Conservative 68; Mismatches 154; Indels 51; Gaps 13;

Qy 2 DEAIAIRAIPIQKWRHQAAREM--ORRCWQIQFQNLEYEASBQDQABLYKPFNDLIKH 58  
 Db 113 DPAATKALITCDPFRIFERFRKAIGGAENEAKISLSCOTLNNSFDANADLANY---- 165

Qy 59 MPOAGRKHQYQGSAAHVSVLDDKODLVVEFGDIDVNAKEIPLIRKHNIDLLDVFRKRGN 118  
 Db 166 ---EGPKLEFE---OLYDDKN-----AFKGAKIKKMSQEFISKMNDFLK-GK 207

Qy 119 RLHPKTYVALIREAAKSILKQLPNISPV--STAVSQQTVCQGDIAHGKULLDLVHLHNGLP 176  
 Db 208 YLPKKTYVAAILISHADTFRQPSPMVLENNTPDVKISVGDTGQFDVNLFRKGKV 267

Qy 177 SSSEPVYENGDFDRGKGKGLBVLLSLYLAPPNAVFLRNTRGHEDSVMMARYGRREV 236  
 Db 268 GPKHTYLFNGDFVDRGSWSCEAVALLFYCLIKLPPNIFLRNKHESDNMKTYGFEDCK 327

Qy 237 SKYPRNRHLILAFIDEVVRMLPIGSVLSRSVLTUHGR-SD-STSDLIKSDRGKVSI 294  
 Db 328 VKY--SORFINNMAQSFESPLATLNLINNDYLMWHGGILPSDPSATLSDPKNIDR----F 379

Qy 295 LRPPLTDGEPPLDKTEWQOQIFDIMSDPOATMGCVENTLIRGAGWMFGBDVTUNFLQRHRLS 354  
 Db 380 AQGP-RDG-----AFMELIWADPQEBANGMGPQO-RGLGHAFDITDRFLRINKUR 428

Qy 355 YVIRSHBLKPGNTHFMHDNKITITFSASNYAIGSNKGAYIRL 397  
 Db 429 KITFRSHELRMGSGVQEOKGKLMVTFPSAPNCDSQGNLGGVIR 471

XX SQ Sequence 513 AA;

Query Match 13.9%; Score 479.5; DB 7; Length 513;  
 Best Local Similarity 32.3%; Pred. No. 9.3e-36;  
 Matches 130; Conservative 68; Mismatches 154; Indels 51; Gaps 13;

Qy 2 DEAIAIRAIPIQKWRHQAAREM--ORRCWQIQFQNLEYEASBQDQABLYKPFNDLIKH 58  
 Db 113 DPAATKALITCDPFRIFERFRKAIGGAENEAKISLSCOTLNNSFDANADLANY---- 165

Qy 59 MPOAGRKHQYQGSAAHVSVLDDKODLVVEFGDIDVNAKEIPLIRKHNIDLLDVFRKRGN 118  
 Db 166 ---EGPKLEFE---OLYDDKN-----AFKGAKIKKMSQEFISKMNDFLK-GK 207

Qy 119 RLHPKTYVALIREAAKSILKQLPNISPV--STAVSQQTVCQGDIAHGKULLDLVHLHNGLP 176  
 Db 208 YLPKKTYVAAILISHADTFRQPSPMVLENNTPDVKISVGDTGQFDVNLFRKGKV 267

RESULT 24 ADK64862





Db	222	IGPERKLEKIVQFPEAVPH	238
RESULT	28		
ID	ABP98033	standard; protein; 323 AA.	
AC	ABP98033;		
XX			
DT	11-AUG-2003	(first entry)	
XX			
DE	Protein phosphatase stress-related polypeptide GmPP2A-4.		
XX			
KW	Protein phosphatase stress-related polypeptide; PPSRP; PPPP2A-1; PPPP-1; BnPP2A-1; BnPP2A-2; BnPP2A-3; GmPP2A-1; GmPP2A-2; GmPP2A-3; OsPP2A-1; OsPP2A-2; OsPP2A-3; plant tolerance; environmental stress; plant; drought; salinity; cold; enzyme.		
XX			
OS	Glycine max.		
XX			
PN	W0003020914-A2.		
XX			
PD	13-MAR-2003.		
XX			
PF	05-SEP-2002; 2002WO-US0208445.		
PR	05-SEP-2001; 2001US-0317305P.		
XX			
PA	(BADI ) BASF PLANT SCI GMBH.		
XX			
PI	Bohnert RJ, Chen R, Ishitani M, Van Thiel N, Da Costa E SilvaO; DR		
XX			
PT	WPI; 2003-300886/29.		
DR	N-PSDB; ACC43348.		
XX			
PT	New protein phosphatase stress-related polypeptide coding nucleic acid, useful for modulating plant's tolerance to an environmental stress such as drought, increased salinity and cold.		
PT	Claim 13; Page 81; 107pp; English.		
XX			
CC	The present sequence represents a protein phosphatase stress-related polypeptide (PPSRP). The specification describes PPSRP polypeptides designated PppP2A-1, PppP2A-1, BnPP2A-3, GmPP2A-1, GmPP2A-2, GmPP2A-3, OsPP2A-1, OsPP2A-2 or OsPP2A-3, derived from Phycomyctrella patens, Brassica napus, Glycine max or Oryza sativa. PPSRP polypeptides and polynucleotides are useful for modulating plant tolerance to an environmental stress such as drought or increased salinity and cold. They are also useful in identification and localization of Phycomyctrella patens, Brassica napus, Glycine max or Oryza sativa and related organisms, mapping or genomes of organisms related to the above species, in a evolutionary and polypeptide structural studies, in determination of PPSRP regions required for function, modulation of PPSRP activity, modulation of metabolism of one or more cell functions and transmembrane transport of one or more components.		
XX	Sequence 323 AA;		
SQ	Query Match	12.7%; Score 435.5; DB 6; Length 323;	
Best Local Similarity	31.2%; Pred. No. 6.4e-32;		
Matches	104; Conservative 66; Mismatches 124; Indels 39; Gaps 9;		
OY	90 DIVMKIELP-IRKHNDLIDVFRKKGRGHLRPPKVALLTREAKSFLKOLPNPSPVSTA 148		
Db	8 DILMLIEBLVTPAKQVQLSSEBIO-----LCVWSREI---FLOORPLNLKLEAP 54		
OY	149 VSQQVTVCQDPLHGKDDLVLUHKQKLSSNPYFNGDFUDRGGRGLEVLILLISLYLA 208		
Db	55 IK---ICGDVHGGQSDLLRLPEYSCGLPPEAN-YFLGLDVYDGRKGSLTCILLYAKIK 109		
209	RPNAPFLNRGNHEDSUMMARYGFIREVEVESKIPRNKRKILAPIDETYRWLPLGSVMSRVL 268		
Db	110 YPENPFKLRGHNCASINRIGYFYDECKRR--NVRLMWTFPDCFCNCLPVAALVDEKIL 166		
OY	269 IWHGGE-S-DTSIDLKSIDRSKVKVSLRPPUTDGERPLDKTEWQFIDMWSPOATMC 327		
Db	167 CMFGGSLSPDLNNLDORINTLQR-----PDPVPDGLCILWNSLPKDVG 211		
OY	328 VENTLRLAGGVWQGPDVTDNFQHRLSVTRSHBECKENGHEFMDNKLTIFASNYTAI 387		
Db	212 WGMNDRGVSYITRCADKVSQFLQKHDILIVCRAHQVEDGYEFFANQLVTFSAPNYCCE 271		
OY	388 GSNKGAYVIRLNQMLMPHRYVQYSAASOTKRLSP 420		
Db	272 FDNAGAMMSVDFIMCSP-QTLKPADKAKLNP 303		
RESULT	29		
ID	ADX87721 standard; protein; 339 AA.		
XX			
AC	ADX87721;		
XX			
DT	21-APR-2005 (first entry)		
XX			
DE	Plant full length insert polypeptide seqid 50385.		
XX			
PR	plant protectant; plant growth regulant; gene therapy; plant; recombinant DNA construct; physical array; plant breeding marker; cold tolerance; heat tolerance; drought tolerance; herbicide tolerance; extreme osmotic condition; pathogen tolerance; pest tolerance; growth rate; cell cycle pathway; disease resistance; galactomannan production; lignin production; Plant growth regulator; yield; plant growth; plant development; seed oil; protein yield; protein content.		
XX			
OS	Unidentified.		
XX			
PN	US2004034888-A1.		
XX			
PD	19-FEB-2004.		
XX			
PR	06-MAY-1999; 99US-00304517.		
PR	05-NOV-2001; 2001US-00985678.		
XX			
PA	(LIUJ/) LIU J.		
PA	(ZHOU/) ZHOU Y.		
PA	(KOVAL/) KOVALIC D K.		
PA	(SCRB/) SCREEN S B.		
PA	(TABR/) TABASKA J E.		
PA	(CAOY/) CAO Y.		
XX			
PI	Liu J, Zhou Y, Kovacic DK, Screen SB, Tabaska JE, Cao Y;		
XX			
DR	WPI; 2004-18013/17.		
XX			
PT	New recombinant DNA construct, useful for improving plant tolerance to cold, heat, drought, herbicides, extreme osmotic conditions, pathogens or pests, for conferring increased resistance to plant disease, or for improving yield.		
XX	Claim 1; SEQ ID NO 50385; 15pp; English.		
XX			
CC	The invention describes a recombinant DNA construct comprising a polynucleotide consisting of a sequence encoding an amino acid sequence available in electronic form from the US patent office at ftp://seqdata.uspto.gov/sequence.html?docid=2004034888. The polynucleotide of the invention are also useful in physical arrays of molecules and as plant breeding markers. The recombinant DNA construct is useful for improving plant tolerance to cold, heat, drought, herbicides, extreme osmotic conditions, pathogens or pests, for manipulating growth rate in plant cells by modification of the cell cycle pathway, for conferring		

CC increased resistance to plant disease, for producing galactomannan, lignin or plant growth regulators, for increasing the rate of homologous recombination in plants, for improving yield by modification of photosynthesis or carbohydrate, nitrogen or phosphorus use and/or uptake or by providing improved plant growth and development under at least one stress condition or for modifying seed oil or protein yield and/or content. This is the amino acid sequence of a plant full length insert polypeptide that can be used in the recombinant DNA construct of the invention.

CC invention.

XX Sequence 339 AA;

SQ Query Match 12.6%; Score 434.5; DB 8; Length 339;

Best Local Similarity 30.1%; Pred. No. 8; Te-32; Mismatches 106; Conservative 72; MisMatches 135; Indels 39; Gaps 9;

OY 56 IKENPQ-AAGRKNQYQCSAHHVSLDDKDLVERFGDVTNAKELPIRKHNIDLLDVFRK 114

Db 1 LAKTPRSPGRK--EGGKAMMTR----APMGPMGAVADEMVR----LVEGGRG 46

OY 115 KQRNRLHPKVYALILRBAKSILKOLPNISPVSTAVSPOCVTVEGDLKGLDPLLVHKNG 174

Db 47 GROVOLSEBAAEIRCOLCVBGKRVILSQPNL---LRHAPVKLGDIHQVPLRLFDGG 102

OY 175 LQSSSNPVPFNGFVDRGKGRGEFLVLLSLVLAFFNAPVFLARGNHEDSVWARYKPIR 234

Db 103 YPPAST-TVPLGDYDVKRQSLKETICLIAKLYKRIPENITFLRGNHEDAKINRYKPYDE 161

OY 235 VESKYPRNHHKRLAFIDEVYRMPLGSVLNSRVLIVHGFS-DSTSIDLISIDRGKVVS 293

Db 162 CRRRF--NVRMKIFSCFCNCPIAALIDKILCMGIGLSBELTSIDQINDIER--- 213

OY 294 ILRPLTDGEPRLDKTEWQIFDIMSDFQATMACCVPNTLIRAGUMFQPDVTFNLFORHL 353

Db 214 -----PARIPDGGLCIDLWSDSPDGPQGWGSDRGVSCTFGAOKVLFBLKNDL 263

OY 354 SVVIRSHCKPNHBFMDNKKITPSASNYAIGSNGKAVIRNLNOLMPHP 405

Db 264 DLICRAHQWVQEDYEFQAQRLLTVTAPSAPNVCGBFDFNVGALLSIDESLMCSP 315

## RESULT 30

AAG30062 ID AAG30062 standard; protein; 318 AA.

XX AC AAG30062;

XX DT 17-OCT-2000 (first entry)

XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 35873.

KW Protein identification; signal transduction pathway; metabolic pathway; hybridisation assay; genetic mapping; gene expression control; promoter; termination sequence.

KW Arabidopsis thaliana.

OS Arabidopsis thaliana.

XX PR P1033405-A2.

XX PD 06-SEP-2000.

XX PR 25-FEB-2000; 2000BP-00301439.

XX PR 25-FEB-1999; 99US-0121825P.

PR 05-MAR-1999; 99US-0123180P.

PR 09-MAR-1999; 99US-0123548P.

PR 23-MAR-1999; 99US-0125788P.

PR 25-MAR-1999; 99US-0126268P.

PR 29-MAR-1999; 99US-0126785P.

PR 01-APR-1999; 99US-0127422P.

PR 06-APR-1999; 99US-0128234P.

PR 08-APR-1999; 99US-0128714P.

PR 16-APR-1999; 99US-0129845P.

PR 19-JUL-1999; 99US-01344334P.

PR 19-APR-1999; 99US-0130077P.

PR 21-APR-1999; 99US-0130499P.

PR 23-APR-1999; 99US-0130510P.

PR 23-APR-1999; 99US-0130591P.

PR 28-APR-1999; 99US-0134499P.

PR 30-APR-1999; 99US-0132048P.

PR 04-MAY-1999; 99US-0134849P.

PR 05-MAY-1999; 99US-013485P.

PR 06-MAY-1999; 99US-0132486P.

PR 07-MAY-1999; 99US-0132863P.

PR 11-MAY-1999; 99US-0134256P.

PR 14-MAY-1999; 99US-0134118P.

PR 14-MAY-1999; 99US-0134212P.

PR 14-MAY-1999; 99US-0134770P.

PR 18-MAY-1999; 99US-0134768P.

PR 19-MAY-1999; 99US-0136941P.

PR 20-MAY-1999; 99US-0135124P.

PR 21-MAY-1999; 99US-0135335P.

PR 24-MAY-1999; 99US-0135629P.

PR 25-MAY-1999; 99US-0136021P.

PR 08-JUN-1999; 99US-0136392P.

PR 28-MAY-1999; 99US-0136782P.

PR 01-JUN-1999; 99US-013722P.

PR 03-JUN-1999; 99US-0137528P.

PR 04-JUN-1999; 99US-0137502P.

PR 07-JUN-1999; 99US-0137724P.

PR 10-JUN-1999; 99US-0138540P.

PR 10-JUN-1999; 99US-0138547P.

PR 14-JUN-1999; 99US-0139119P.

PR 16-JUN-1999; 99US-0139452P.

PR 16-JUN-1999; 99US-0139453P.

PR 17-JUN-1999; 99US-0139492P.

PR 18-JUN-1999; 99US-0139454P.

PR 18-JUN-1999; 99US-0139455P.

PR 18-JUN-1999; 99US-0139456P.

PR 18-JUN-1999; 99US-0139457P.

PR 18-JUN-1999; 99US-0139458P.

PR 18-JUN-1999; 99US-0139459P.

PR 18-JUN-1999; 99US-0139460P.

PR 18-JUN-1999; 99US-0139461P.

PR 18-JUN-1999; 99US-0139462P.

PR 18-JUN-1999; 99US-0139463P.

PR 18-JUN-1999; 99US-0139464P.

PR 18-JUN-1999; 99US-0139465P.

PR 18-JUN-1999; 99US-0139466P.

PR 18-JUN-1999; 99US-0139467P.

PR 18-JUN-1999; 99US-0139468P.

PR 18-JUN-1999; 99US-0139469P.

PR 18-JUN-1999; 99US-0139470P.

PR 18-JUN-1999; 99US-0139471P.

PR 18-JUN-1999; 99US-0139472P.

PR 18-JUN-1999; 99US-0139473P.

PR 18-JUN-1999; 99US-0139474P.

PR	19-JUL-1999;	99US-0144335P.
PR	20-JUL-1999;	99US-0144322P.
PR	20-JUL-1999;	99US-0144884P.
PR	21-JUL-1999;	99US-0144844P.
PR	21-JUL-1999;	99US-0145086P.
PR	21-JUL-1999;	99US-0145086P.
PR	22-JUL-1999;	99US-0145085P.
PR	22-JUL-1999;	99US-0145087P.
PR	22-JUL-1999;	99US-0145088P.
PR	22-JUL-1999;	99US-0145122P.
PR	23-JUL-1999;	99US-0145145P.
PR	23-JUL-1999;	99US-0145218P.
PR	23-JUL-1999;	99US-0145244P.
PR	26-JUL-1999;	99US-0145265P.
PR	27-JUL-1999;	99US-0145913P.
PR	27-JUL-1999;	99US-0145918P.
PR	28-JUL-1999;	99US-0145919P.
PR	02-AUG-1999;	99US-0146316P.
PR	02-AUG-1999;	99US-0146388P.
PR	02-AUG-1999;	99US-0146389P.
PR	03-AUG-1999;	99US-0147038P.
PR	04-AUG-1999;	99US-0147204P.
PR	04-AUG-1999;	99US-0147302P.
PR	05-AUG-1999;	99US-0147322P.
PR	06-AUG-1999;	99US-0147360P.
PR	06-AUG-1999;	99US-0147416P.
PR	09-AUG-1999;	99US-0147493P.
PR	09-AUG-1999;	99US-0147935P.
PR	10-AUG-1999;	99US-0148171P.
PR	11-AUG-1999;	99US-0148199P.
PR	12-AUG-1999;	99US-0148341P.
PR	13-AUG-1999;	99US-014855P.
PR	13-AUG-1999;	99US-014664P.
PR	16-AUG-1999;	99US-014968P.
PR	17-AUG-1999;	99US-014975P.
PR	18-AUG-1999;	99US-014926P.
PR	20-AUG-1999;	99US-014972P.
PR	20-AUG-1999;	99US-014923P.
PR	20-AUG-1999;	99US-014929P.
PR	23-AUG-1999;	99US-014992P.
PR	23-AUG-1999;	99US-0149930P.
PR	25-AUG-1999;	99US-0150566P.
PR	26-AUG-1999;	99US-0150894P.
PR	27-AUG-1999;	99US-015105P.
PR	27-AUG-1999;	99US-015106P.
PR	27-AUG-1999;	99US-015108P.
PR	30-AUG-1999;	99US-0151303P.
PR	31-AUG-1999;	99US-015138P.
PR	07-SEP-1999;	99US-015233P.
PR	10-SEP-1999;	99US-0153070P.
PR	13-SEP-1999;	99US-0153158P.
PR	15-SEP-1999;	99US-0154018P.
PR	16-SEP-1999;	99US-0154039P.
PR	20-SEP-1999;	99US-0154779P.
PR	04-OCT-1999;	99US-0157117P.
PR	22-SEP-1999;	99US-0155139P.
PR	23-SEP-1999;	99US-0155486P.
PR	24-SEP-1999;	99US-0155559P.
PR	28-SEP-1999;	99US-015658P.
PR	29-SEP-1999;	99US-015696P.
PR	04-OCT-1999;	99US-0157717P.
PR	05-OCT-1999;	99US-0157753P.
PR	06-OCT-1999;	99US-015865P.
PR	07-OCT-1999;	99US-0158039P.
PR	08-OCT-1999;	99US-0158232P.
PR	12-OCT-1999;	99US-0158399P.
PR	13-OCT-1999;	99US-0159233P.
PR	13-OCT-1999;	99US-0159252P.
PR	14-OCT-1999;	99US-0159339P.
PR	14-OCT-1999;	99US-0159331P.
PR	14-OCT-1999;	99US-0159337P.
PR	14-OCT-1999;	99US-0159338P.
PR	18-OCT-1999;	99US-0159584P.
PR	21-OCT-1999;	99US-0160741P.
PR	21-OCT-1999;	99US-0160767P.
PR	21-OCT-1999;	99US-0160768P.
PR	21-OCT-1999;	99US-0160770P.
PR	21-OCT-1999;	99US-0160814P.
PR	21-OCT-1999;	99US-0160815P.
PR	22-OCT-1999;	99US-0160880P.
PR	22-OCT-1999;	99US-0160881P.
PR	22-OCT-1999;	99US-0160899P.
PR	25-OCT-1999;	99US-0161340P.
PR	25-OCT-1999;	99US-0161405P.
PR	25-OCT-1999;	99US-0161406P.
PR	26-OCT-1999;	99US-0161359P.
PR	26-OCT-1999;	99US-0161360P.
PR	26-OCT-1999;	99US-0161361P.
PR	28-OCT-1999;	99US-0161920P.
PR	28-OCT-1999;	99US-0161922P.
PR	28-OCT-1999;	99US-0161939P.
PR	29-OCT-1999;	99US-0162142P.
QY	59	MEDQAAGERKNQYQGSAHVSUDDKDLVEEVNATCDEWVAKIEPLRKNHIDLLDVFRRKR--
Db	1	MARPKPAQEQBQEPRAVLD-----DIRRLV-----FANTRPG : 116
QY	117	-GNRLHPKYVALLRRAAKSLKQLPNTS-----PVASTAVQQVTVCGLDHGKLDDLV 168
Db	37	SKQKQH-----LSBG--EIRQLCAVSKETIPLQQPNLLEBAPIKICGDIHQISDLIR 87
QY	169	VIAHKQNLPPSSNPYVNGDPDRGKGEVLILLILSYLVAFFNPAFLNRGHEDSNVAR 228
Db	147	VGFYDECKRKKF---NVRWKIFTDCNCNCLPVALIDRILCHGGISPEJKSLDQIRNA 203
QY	288	RKGYVSLVIRPPLTDGSPDLKTVWQOIFDMDSPDTMGCVPNTLRGAGYMGFGPDTNF 347
Db	204	R-----PMDIPESGLVCDLLWSDPSGDVGMND-RGVSVTFFGADKVAF 247
QY	348	LQRHLISVYVRSHECKENGHRPMHDNKITIIFASNYVIAIGNSKGAYIRLANQLMRPH 405
Db	248	LEKHDMDLICRAHQVVEDGYEFABRQLVTVFSAFPYCGBFDNAGAMMSIDESLWCSF 305
RESULT	31	ADN72465
ID	ADN72465	standard; protein; 318 AA.
AC	ADN72465;	
XX		
DT	15-JUL-2004	(first entry)
XX		
DE	Thale cress protein upregulated in B2Fa/Dpa expressing plants	SqId 360.
XX		
KW	Plant; transgenic; B2Fa/Dpa transcription factor; growth regulator;	
KW	animal feed product; thale cress; cell wall biosynthesis;	
KW	nitrogen metabolism; carbon metabolism.	
XX		
OS	Arabidopsis thaliana.	
XX		
PN	WO2004035798-A2.	
XX		
PD	29-APR-2004.	



Db 62 DIINRILLEV-----RSRPKGKVQVLSBSBIRHLCAASRETFIPLQPNLLELA 107  
 Qy :|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:  
 Db 148 AVSQQTVCQDGLKGKDDLLJVLWHLNGKLSSNPVFGNDPDRKGKGLEVLLLISLY 207  
 Qy :|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:  
 Db 108 PIK----ICDPVHGCGSPDRKPLPEYGGLPFPEAN-TLFGLGVDQKQSLTICLAVKL 162  
 Qy :|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:  
 208 APPNAVFLNKGHEHDSVMARYGTRVEVKSPRNHRKIAFIDEVRYMPVPLGSVLSRV 267  
 163 KYPENPFULLKQHNECASINRYGFIDECKERP--NVRWKTKTFECFCNCLPVAAVIDEKI 219  
 Qy :|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:  
 Db 268 LIVSGGFS-DTSLDLKSIDRKGKVSLIRPLPLDGRPKTKHQPIQIPIMWDQATNG 326  
 Qy :|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:  
 Db 220 LCMHGGGLSPDILNLQDQRNR-----PTDVPPDGLCDLWSDPSKEVQ 264  
 327 CVPNTLARGAGWGRGPDTMFLQRHLSTVIRSHCKPNCHEFDNKLTITASNYTA 386  
 265 GWMQMDRGVSYTGAKVSKFLQKDLICRAHQVWBGYRFPANQVTISAPNCG 324  
 Qy :|||:|||:|||:|||:|||:|||:|||:|||:|||:  
 387 IGSNKGAYTIRLNLQMLPHFVQYISASQTCRLSF 420  
 Db 325 EFDNAGAMMSVDETLMCSF-QILKPADKKAKLNF 357

## RESULT 33

ADT59840 ID ADT59840 standard; protein: 325 AA.

AC XX ADT59840;  
 XX DT 13-JAN-2005 (first entry)

DB Plant polypeptide, SEQ ID 9917.

XX Plant; transgenic; cold tolerance; growth rate; drought tolerance;

KW disease resistance; galactomannan production; plant growth regulator; heat tolerance; herbicide tolerance; lignin production; extreme osmotic condition tolerance; pathogen resistance; pest resistance; yield improvement; seed oil yield; seed protein yield. Viridiplantae.

XX OS US2004216190-A1.

XX PD 28-OCT-2004.

XX PP 18-DEC-2003; 2003US-00739930.

XX PR 28-APR-2003; 2003US-00424599.

XX PR 28-APR-2003; 2003US-00425115.

XX PA (KOVA/) KOVALIC D K.

XX PI Kovalic DK;

XX DR WPI; 2004-757369/14.

XX New recombinant DNA constructs useful in the field of biochemistry and genetics, and in particular for producing transgenic plants with improved biological characteristics.

XX PS Claim 2; SEQ ID NO 9917; 14bp; English.

The invention relates a recombinant DNA construct comprising a polynucleotide having any of 5544 nucleotide sequences (cDNAs) SEQ ID NO: 1-5544) and encoding a polypeptide with any of 5544 amino acid sequences (SEQ ID NO: 544-1108). The cDNAs and proteins are from corn, soybean, Arabidopsis, wheat and rape but the specification does not indicate which sequences is derived from which organism. Also included is a method of producing a plant having an improved property, comprising transforming a plant with a recombinant DNA construct comprising a promoter region functional in a plant cell operably joined to a polynucleotide encoding a polypeptide associated with the property, and growing the transformed plant. The property is selected from improving plant cold tolerance, for

CC manipulating growth rate in plant cells by modification of the cell cycle pathway, for improving plant drought tolerance, for providing increased resistance to plant disease, for galactomannan production, for production of plant growth regulators, for improving plant heat tolerance, for improving plant tolerance to extreme osmotic conditions, for improving plant photosynthesis, for modifying seed oil yield and/or content, for modifying seed protein Yield and/or content, for yield improvement by modification of carbohydrate, nitrogen or phosphorus use and/or uptake and for yield improvement by providing improved plant growth and development under at least one stress condition. The polynucleotide may also encode a plant transcription factor. The methods and compositions of the present invention are useful in the field of biochemistry and genetics, in particular for producing transgenic plants with improved biological characteristic such as increased yield, improved nitrogen flow, increasing plant tolerance to cold or heat, improving plant tolerance to extreme osmotic and drought conditions, and improving plant tolerance to plant pests or pathogens. They can also be used in physical arrays of molecules, plant breeding markers, computer-based storage and analysis systems. The present sequence is one of the 5544 plant protein sequences of the invention. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from USPTO at seqdata.uspto.gov/sequence.html?docID=20040216190.

## Sequence 325 AA:

Query Match 12.5%; Score 430.5; DB 8; Length 325;  
 Best Local Similarity 31.1%; Pred. No. 1\_9e-31;  
 Matches 99; Conservative 68; Mismatches 122; Indels 29; Gaps 6;

Qy 89 GDTVNKAELBLPIRKNHIDLIVFLYFRKRKGNNLRPKVAVLIRRAKSUQLQPLNISPVSTA 148  
 Db 9 GAMEGALDLSVVR---LVEGRGGRQVOVSBEAIKLCVDAKRVPLSOPNL---LR 59

Qy 149 VSQQTVCQDGLKGKDDLLJVLWHLNGKLSSNPVFGNDPDRKGKGLEVLLLISLYA 208  
 Db 60 IQARPKCQDIDQPVDRPLPFLGGVPRST-YVFLPDYDVGKQSLTICLAVKL 118

Qy 209 FPMAVFLNKGHEHDSVMARYGTRVEVKSPRNHRKIAFIDEVRYMPVPLGSVLSRV 268  
 Db 119 YDKVPLFLRGNHDAKIRVYGFIDECKERP--NVRWKTKTFECFCNCLPVAAVIDEKI 175

Qy 269 IVSGGFS-DTSLDLKSIDRKGKVSLIRPLPLDGRPKTKHQPIQIPIMWDQATNGC 327

Db 176 CMHGGSLSPDILNLQDQR-----PTDVPPDGLCDLWSDPSKEVQ 220

Qy 328 VPKVLRGAGWGRGPDTMFLQRHLSTVIRSHCKPNCHEFDNKLTITASNYTA 387  
 Db 221 WGSDRGVSYTGAKVSKFLQKDLICRAHQVWBGYRFPANQVTISAPNCG 280

Qy 388 GSNKGAYTIRLNLQMLPHFVQYISASQTCRLSF 405  
 Db 281 FDNVGALLSIDENLMCSF 298

## RESULT 34

ADT08660 ID ADT08660 standard; protein: 334 AA.

AC XX ADT08660;  
 AC DT 21-APR-2005 (first entry)

XX DB Plant full length insert polypeptide seqid 64475.

XX plant protectant; plant growth regulator; gene therapy; plant; recombinant DNA construct; physical array; plant breeding marker; cold tolerance; heat tolerance; drought tolerance; herbicide tolerance; extreme osmotic condition; pathogen tolerance; pest tolerance; growth rate; cell cycle pathway; disease resistance;





PR 19-JUL-1999; 99US-0144333P.  
 PR 19-JUL-1999; 99US-0144334P.  
 PR 19-JUL-1999; 99US-0144335P.  
 PR 20-JUL-1999; 99US-0144322P.  
 PR 20-JUL-1999; 99US-0144622P.  
 PR 21-JUL-1999; 99US-0144884P.  
 PR 21-JUL-1999; 99US-0145066P.  
 PR 21-JUL-1999; 99US-0145080P.  
 PR 22-JUL-1999; 99US-0145085P.  
 PR 22-JUL-1999; 99US-0145087P.  
 PR 22-JUL-1999; 99US-0145089P.  
 PR 22-JUL-1999; 99US-0145122P.  
 PR 23-JUL-1999; 99US-014515P.  
 PR 23-JUL-1999; 99US-0145218P.  
 PR 23-JUL-1999; 99US-0145224P.  
 PR 26-JUL-1999; 99US-0145266P.  
 PR 27-JUL-1999; 99US-0145913P.  
 PR 27-JUL-1999; 99US-0145918P.  
 PR 27-JUL-1999; 99US-0145919P.  
 PR 28-JUL-1999; 99US-0145931P.  
 PR 02-AUG-1999; 99US-0146366P.  
 PR 02-AUG-1999; 99US-0146389P.  
 PR 02-AUG-1999; 99US-0146399P.  
 PR 03-AUG-1999; 99US-0147038P.  
 PR 04-AUG-1999; 99US-0147202P.  
 PR 05-AUG-1999; 99US-0147202P.  
 PR 06-AUG-1999; 99US-0147303P.  
 PR 09-AUG-1999; 99US-0147493P.  
 PR 10-AUG-1999; 99US-0148171P.  
 PR 11-AUG-1999; 99US-0148319P.  
 PR 12-AUG-1999; 99US-0148319P.  
 PR 13-AUG-1999; 99US-014855P.  
 PR 16-AUG-1999; 99US-014864P.  
 PR 17-AUG-1999; 99US-0149175P.  
 PR 18-AUG-1999; 99US-0149426P.  
 PR 20-AUG-1999; 99US-0149722P.  
 PR 20-AUG-1999; 99US-0149723P.  
 PR 23-AUG-1999; 99US-0149902P.  
 PR 23-AUG-1999; 99US-0149930P.  
 PR 25-AUG-1999; 99US-0150566P.  
 PR 26-AUG-1999; 99US-015084P.  
 PR 27-AUG-1999; 99US-0151066P.  
 PR 27-AUG-1999; 99US-0151080P.  
 PR 30-AUG-1999; 99US-0151303P.  
 PR 31-AUG-1999; 99US-0151438P.  
 PR 01-SEPT-1999; 99US-0151930P.  
 PR 07-SEPT-1999; 99US-015263P.  
 PR 10-SEPT-1999; 99US-0153070P.  
 PR 13-SEPT-1999; 99US-0153758P.  
 PR 15-SEPT-1999; 99US-0154018P.  
 PR 16-SEPT-1999; 99US-0154039P.  
 PR 20-SEPT-1999; 99US-0154779P.  
 PR 22-SEPT-1999; 99US-0155139P.  
 PR 23-SEPT-1999; 99US-015586P.  
 PR 24-SEPT-1999; 99US-015589P.  
 PR 28-SEPT-1999; 99US-0156458P.  
 PR 29-SEPT-1999; 99US-015656P.  
 PR 04-OCT-1999; 99US-0157117P.  
 PR 05-OCT-1999; 99US-0157753P.  
 PR 06-OCT-1999; 99US-0157865P.  
 PR 07-OCT-1999; 99US-0158039P.  
 PR 08-OCT-1999; 99US-0158332P.  
 PR 12-OCT-1999; 99US-0158569P.  
 PR 13-OCT-1999; 99US-0159233P.  
 PR 13-OCT-1999; 99US-015935P.

PR 13-OCT-1999; 99US-015935P.  
 PR 14-OCT-1999; 99US-0159329P.  
 PR 14-OCT-1999; 99US-0159330P.  
 PR 14-OCT-1999; 99US-0159331P.  
 PR 14-OCT-1999; 99US-0159337P.  
 PR 14-OCT-1999; 99US-0159338P.  
 PR 18-OCT-1999; 99US-015984P.  
 PR 21-OCT-1999; 99US-016041P.  
 PR 21-OCT-1999; 99US-0160767P.  
 PR 21-OCT-1999; 99US-0160788P.  
 PR 21-OCT-1999; 99US-0160770P.  
 PR 21-OCT-1999; 99US-0160814P.  
 PR 21-OCT-1999; 99US-0160915P.  
 PR 22-OCT-1999; 99US-0160880P.  
 PR 22-OCT-1999; 99US-0160981P.  
 PR 22-OCT-1999; 99US-0160989P.  
 PR 25-OCT-1999; 99US-0161404P.  
 PR 25-OCT-1999; 99US-0161405P.  
 PR 25-OCT-1999; 99US-0161406P.  
 PR 26-OCT-1999; 99US-0161359P.  
 PR 26-OCT-1999; 99US-0161360P.  
 PR 26-OCT-1999; 99US-0161361P.  
 PR 28-OCT-1999; 99US-0161920P.  
 PR 28-OCT-1999; 99US-0161920P.  
 PR 28-OCT-1999; 99US-0161933P.  
 PR 29-OCT-1999; 99US-0162142P.

Query Match 12.5%; Score 430; DB 3; Length 304;  
 Best Local Similarity 31.2%; Pred. No. 1.9e-31;  
 Matches 102; Conservative 61; Mismatches 110; Indels 54; Gaps 10;

QY 90 DIVNAKELPLPKRNHIDLIDVFRKR--GNRLAPRKYVALILREANKSKLQLPNIS--- 143  
 ||| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |  
 Db 8 DIRRIV-E-----FRNTRPGSOKQVH----LSEG--EIRQLCAVSKIEI 44  
 QY 144 ----PVSTAVSQVTYCGDGLQKLDDLVLAHKNGLPPSSNPYVENGDFVDRGKGRLVL 199  
 ||| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |  
 Db 45 FLQQPNMELLEPLAKIIGDGHDQYSDLRFLPEYGGFPPEAN-YLFEGDYVDRGKOSLI 103  
 QY 200 LLLSILYLAFFNAVPLRGRNHEDSVWAMAROFIREEVSKPRHNRKLLAFDDEVYRMPL 259  
 ||| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |  
 Db 104 CLLAXXKIKYENFELLRGNTESASINRITYGFDCKRRP--NTRLWKFIDCENCILWV 160  
 ||| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |  
 QY 260 GSVLNSRVLVHGGS-DSTSLLDKSIDRGKVYSLTRPLTDGSPLDKTEWQOIFDMM 318  
 :::: | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |  
 Db 161 AALIDDRILCMHFGGJSPELKSLDQJRNIA-----PMIDIPESGLVCDLW 205  
 ||| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |  
 QY 319 SDPOATMGCVNTLQAGAAGVFGDPDVMDNLQRHLSYVTSHECKPNGHFRPHDKITI 378  
 ||| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |  
 Db 206 SDPSGDVGWMD-RGSVYTFGADKVAEFLERKHDMLICRAHQVVBGYBPFABROLTV 264  
 ||| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |  
 QY 379 FSASNTYAIIGSNKGAYIIRUNQOLMPH 405  
 ||| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |  
 Db 265 FSAPNYCGKEENAGAMMSISLMSF 291  
 ||| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |  
 RESULT 37  
 ADT57338  
 ID ADT57338 standard; protein; 325 AA.  
 AC ADT57338;  
 XX  
 DT 13-JAN-2005 (first entry)  
 DB  
 Plant polypeptide, SEQ ID 7415.  
 XX  
 Plant; transgenic; cold tolerance; growth rate; drought tolerance;  
 KW disease resistance; galactomannan production; plant growth regulator;  
 KW heat tolerance; herbicide tolerance; lignin production;  
 KW extreme isomeric condition tolerance; pathogen resistance;  
 KW pest resistance; yield improvement; seed oil yield; seed protein yield.  
 XX  
 OS Viridiplantae.



homologous recombination in plants, for lignin production, for improving plant tolerance to extreme osmotic conditions, for improving plant tolerance to pathogens or pests, for yield improvement by modification of photosynthesis, for modifying seed oil yield and/or content, for modifying seed protein yield and/or content, for yield improvement by modification of carbohydrate, nitrogen or phosphorus use and/or uptake and for yield improvement by providing improved plant growth and development under at least one stress condition. The polynucleotide may also encode a plant transcription factor. The methods and compositions of the present invention are useful in the field of biochemistry and genetics, in particular for producing transgenic plants with improved biological characteristics such as increased yield, improved nitrogen flow, increasing plant tolerance to cold or heat, improving plant tolerance to extreme osmotic and drought conditions, and improving plant tolerance to plant pests or pathogens. They can also be used in physical arrays of molecules, plant breeding markers, computer-based storage and analysis systems. The present sequence is one of the 554 plant protein sequences of the invention. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from USPTO at <http://seqdata.uspto.gov/sequence.html?DocID=20040216190>.

SQ Sequence 326 AA;

Query Match 12.5%; Score 429.5; DB 8; Length 326;

Best Local Similarity 30.5%; Pred. No. 2.4e-31; Matches 102; Conservative 66; Mismatches 125; Indels 41; Gaps 8; PT DR

QY 90 DIVNAKABLPIRKNHIDLLIDVPRKKRGN-RHLPKVVALTREAAKSLSKLQPNLSPVST 147  
Db 8 DIINRLIEV-----RGRGKQVOLSEBIRQLCAASRBFQLOPNLLEBA 53  
QY 148 AVSQQTVCVGDLAKGKLDLILVLUHKNGLPSSENPPVNGDFDGRGRGLSYLWLSYL 207  
Db 54 PIK---ICGDVHGQYSDLLRFLFEGYGGLPPBAN-YLFGLGYDVRKGQSLETICLILAYKI 108  
QY 208 ASPNPAVNLNRGHHEDSYNNARYGFIREVESKPRHRKILLAFIDEYRMILPLGSVNSRV 267  
Db 109 KYPENFFLRLRGHBCASINRTRYGPYDCKRF--NVRWKTFPTECNCLPVAALEDEI 165  
QY 268 LIVHGGES-DSTSLSLDLKSIIDRKYVSLRPLPDTGEPLDKTWEQQFDIMSDPORTMG 326  
Db 166 ICMHGGIS-PDPDLNLIDQTRNLT-----PTDPDTGICLQJLWSDEKEVO 210  
QY 327 CVPNTLRAGAGWFGPDYTDNFQRHRHSYVTSHECKPNKGHRPMHNKINTIFSASNYA 386  
Db 211 GKHMDRKGVSYTFGADKVSFQKHDLICRRAHQVVEDGEKFFPANKQLVITFASPNCG 270  
SQ Sequence 326 AA;

Query Match 12.5%; Score 429; DB 8; Length 338;

Best Local Similarity 31.8%; Pred. No. 2.9e-31; Matches 103; Conservative 60; Mismatches 129; Indels 32; Gaps 7; PT DR

QY 94 AKIELLR-----KNHDLLIVFRKGKGRNLRHLPKYALI--LRAAKSLIKOLPNI 142  
Db 1 AGISLSPRSLSLPGEAMDEBAVDLDIRLLEAGRRT-PRNAQVTDARIRLCAAKDVFLS 59  
QY 143 SPVSTAVSQVTVCGLHKGDDLILVLUHKNGLPSSENPPVNGDFDGRGRGLSYLUL 202  
Db 60 QPNLILRALKPKICGDVHGQYSDLLRFLFEGYGGPDAAN-TLFGLGYDVRGSQISTCIL 118  
SQ Sequence 338 AA;

Query Match 12.5%; Score 429; DB 8; Length 338;

Best Local Similarity 31.8%; Pred. No. 2.9e-31; Matches 103; Conservative 60; Mismatches 129; Indels 32; Gaps 7; PT DR

QY 203 ISLYLAFFPNAVFLNRGHHEDSYNNARYGFIREVESKPRHRKILLAFIDEYRMILPLGSV 262  
Db 119 LAYKIKRPPENFVLRLRGHBCASINRTRYGPYDCKRF--NVRWKTFPTECNCLPVAA 175  
QY 263 LNSRVLVIRSGCSFS-DSTSLSLDLKSIIDRKYVSLRPLPDTGEPLDKTWEQQFDIMSD 321  
Db 176 IDDKLFCMGGGLSPOLSKMDQJLNRFLP-----PVDVDPGLICLWSDP 220  
QY 322 QATMGCGUTLRTAGAGWFGPDYTDNFQRHRHSYVTSHECKPNKGHRPMHNKINTIFS 381  
Db 221 DKRIDRKGVSYTFGADKVSFQKHDLICRRAHQVVEDGEKFFPANKQLVITFASPNCG 280  
OS Unidentified.

homologous recombination in plants, for lignin production, for improving plant tolerance to extreme osmotic conditions, for improving plant tolerance to pathogens or pests, for yield improvement by modification of photosynthesis, for modifying seed oil yield and/or content, for modifying seed protein yield and/or content, for yield improvement by modification of carbohydrate, nitrogen or phosphorus use and/or uptake and for yield improvement by providing improved plant growth and development under at least one stress condition. The polynucleotide may also encode a plant transcription factor. The methods and compositions of the present invention are useful in the field of biochemistry and genetics, in particular for producing transgenic plants with improved biological characteristics such as increased yield, improved nitrogen flow, increasing plant tolerance to cold or heat, improving plant tolerance to extreme osmotic and drought conditions, and improving plant tolerance to plant pests or pathogens. They can also be used in physical arrays of molecules, plant breeding markers, computer-based storage and analysis systems. The present sequence is one of the 554 plant protein sequences of the invention. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from USPTO at <http://seqdata.uspto.gov/sequence.html?DocID=20040216190>.

XN US2004034888-A1.  
PN 06-MAY-1999; 99US-00394517.  
XX 05-NOT-2001; 2001US-00985678.  
PR (ZHOU/) LIU J.  
PA (KOVACI) KOVALIC D.K.  
PA (SCREW) SCREEN S.E.  
PA (TABATA) TABASKA J.B.  
PA (CAO/) CAO Y.

XX DR Liu J, Zhou Y, Kovalic DK, Screen SE, Tabaska JB, Cao Y;  
XX WPI; 2004-180133/17.  
XX PT New recombinant DNA construct, useful for improving plant tolerance to cold, heat, drought, herbicides, extreme osmotic conditions, pathogens or pests, for conferring increased resistance to plant disease, or for improving yield.  
XX PS Claim 1; SEQ ID NO 66618; 15pp; English.

The invention describes a recombinant DNA construct comprising a polynucleotide consisting of a sequence encoding an amino acid sequence available in electronic form from the US patent office at [ftp://seqdata.uspto.gov/sequence.html?DocID=2004034888](http://seqdata.uspto.gov/sequence.html?DocID=2004034888). The polynucleotide of the invention are also useful in physical arrays of molecules and as plant breeding markers. The recombinant DNA construct is useful for improving plant tolerance to cold, heat, drought, herbicides, extreme osmotic conditions, pathogens or pests, for manipulating growth rate in plant cells by modification of the cell cycle pathway, for conferring increased resistance to plant disease, for producing galactomannan, lignin or plant growth regulators, for increasing the rate of homologous recombination in plants, for improving yield by modification of photosynthesis or carbohydrate, nitrogen or phosphorus use and/or uptake or by providing improved plant growth and development under at least one stress condition or for modifying seed oil or protein yield and/or content. This is the amino acid sequence of a plant full length insert polypeptide that can be used in the recombinant DNA construct of the invention.

SQ Sequence 338 AA;

Query Match 12.5%; Score 429; DB 8; Length 338;

Best Local Similarity 31.8%; Pred. No. 2.9e-31; Matches 103; Conservative 60; Mismatches 129; Indels 32; Gaps 7; PT DR

QY 94 AKIELLR-----KNHDLLIVFRKGKGRNLRHLPKYALI--LRAAKSLIKOLPNI 142  
Db 1 AGISLSPRSLSLPGEAMDEBAVDLDIRLLEAGRRT-PRNAQVTDARIRLCAAKDVFLS 59  
QY 143 SPVSTAVSQVTVCGLHKGDDLILVLUHKNGLPSSENPPVNGDFDGRGRGLSYLUL 202  
Db 60 QPNLILRALKPKICGDVHGQYSDLLRFLFEGYGGPDAAN-TLFGLGYDVRGSQISTCIL 118  
SQ Sequence 338 AA;

Query Match 12.5%; Score 429; DB 8; Length 338;

Best Local Similarity 31.8%; Pred. No. 2.9e-31; Matches 103; Conservative 60; Mismatches 129; Indels 32; Gaps 7; PT DR

QY 203 ISLYLAFFPNAVFLNRGHHEDSYNNARYGFIREVESKPRHRKILLAFIDEYRMILPLGSV 262  
Db 119 LAYKIKRPPENFVLRLRGHBCASINRTRYGPYDCKRF--NVRWKTFPTECNCLPVAA 175  
QY 263 LNSRVLVIRSGCSFS-DSTSLSLDLKSIIDRKYVSLRPLPDTGEPLDKTWEQQFDIMSD 321  
Db 176 IDDKLFCMGGGLSPOLSKMDQJLNRFLP-----PVDVDPGLICLWSDP 220  
QY 322 QATMGCGUTLRTAGAGWFGPDYTDNFQRHRHSYVTSHECKPNKGHRPMHNKINTIFS 381  
Db 221 DKRIDRKGVSYTFGADKVSFQKHDLICRRAHQVVEDGEKFFPANKQLVITFASPNCG 280  
OS Unidentified.

Query Match 12.5%; Score 429; DB 8; Length 338;  
 Best Local Similarity 31.8%; Pred. 2.9e-31; Matches 103; Conservative 60; Mismatches 129; Indels 32; Gaps 7;

**RESULT 40**

OY 382 SNNYAIIGSNKAGAYTRINQNLMPHP 405  
 ||| : : | : | : |  
 Db 281 PNYCGBFDNAGALMSIDNSIVCSP 304

**RESULT 40**

ADY06935 ADY06935 standard; protein; 338 AA.  
 XX  
 AC ADY06935;  
 XX DT 21-APR-2005 (first entry)  
 XX DB Plant full length insect polypeptide seqid 62750.  
 XX KW plant protectant; plant growth regulator; gene therapy; plant;  
 KW recombinant DNA construct; physical array; plant breeding marker;  
 KW cold tolerance; heat tolerance; drought tolerance; herbicide tolerance;  
 KW extreme osmotic condition; pathogen tolerance; pest tolerance;  
 KW growth rate; cell cycle pathway; disease resistance;  
 KW galactomannan production; lignin production; plant growth regulator;  
 KW yield; plant growth; plant development; seed oil; protein yield;  
 KW protein content.

XX OS Unidentified.  
 XX US2004034888-A1.  
 XX PD 19-FEB-2004.  
 XX PP 28-APR-2003; 2003US-00425114.  
 XX PR 06-MAY-1999; 99US-00304517.  
 XX PR 05-NOV-2001; 2001US-00985678.

PA (LIU/) LIU J.  
 PA (ZHOU/) ZHOU Y.  
 PA (KOVALIC D K.  
 PA (SCREB) SCREB S E.  
 PA (TABASCA J E.  
 PA (CAOY/) CAO Y.  
 XX PI Liu J, Zhou Y, Kovalic DK, Screen SE, Tabaska JB, Cao Y;  
 XX DR WPI; 2004-180133/17.

PT New recombinant DNA construct, useful for improving plant tolerance to cold, heat, drought, herbicides, extreme osmotic conditions, pathogens or pests, for conferring increased resistance to plant disease, or for improving yield.

XX PS Claim 1; SEQ ID NO 62750; 15pp; English.

XX CC The invention describes a recombinant DNA construct comprising a polynucleotide consisting of a sequence encoding an amino acid sequence available in electronic form from the US patent office at fpc.seqdata.uspto.gov/Sequence.html?pocto\_id=2004034888. The polynucleotide of the invention are also useful in physical arrays of molecules and as plant breeding markers. The recombinant DNA construct is useful for improving plant tolerance to cold, heat, drought, herbicides, extreme osmotic conditions, pathogens or pests, for manipulating growth rate in plant cells by modification of the cell cycle pathway, for conferring increased resistance to plant disease, for producing galactomannin, lignin or plant growth regulators, for increasing the rate of homologous recombination in plants, for improving yield by modification of photosynthesis or carbohydrate, nitrogen or phosphorus use and/or uptake or by providing improved plant growth and development under at least one stress condition or for modifying seed oil or protein yield and/or content. This is the amino acid sequence of a plant full length insect polypeptide that can be used in the recombinant DNA construct of the invention.

XX SQ Sequence 338 AA;

Query Match 12.5%; Score 429; DB 8; Length 338;  
 Best Local Similarity 31.8%; Pred. 2.9e-31; Matches 103; Conservative 60; Mismatches 129; Indels 32; Gaps 7;

**RESULT 40**

OY 94 AKELPLR-----KRHDLILDVFRKCRSRHLRKYVALI--DRAEAKSLKOPNT 142  
 ||| : : | : | : | : | : | : | : | : | : |  
 Db 1 AGISLPSRSPPGRAMDASRAVDLIRRLLBARGGT-PRNAQTDAAEIRRLCAAQDYFLS 59

**RESULT 40**

OY 143 SPVSTASQSVTWCGDPLGKGKDPLVTLHKONGLPSSSNPVSYVNGDFDRGKGLEVLILL 202  
 ||| : : | : | : | : | : | : | : | : | : | : |  
 Db 60 QPNLLELLAPKIKCGDVGQYSDLRLPEYGGYPDAN-VLFITGDDYDRGKOSIETCIL 118  
 ||| : : | : | : | : | : | : | : | : | : | : |  
 OY 203 LSIYLAPFNAVPLRGHEDSMWARMGPIREVESKPRNHGKILADIEBVRLWPLGSV 262  
 ||| : : | : | : | : | : | : | : | : | : | : |  
 Db 119 LAVKIKYPENPFNLRLGRHNECASINRYYGFPDKRR--NTRIKPTECNCLPVPAI 175  
 ||| : : | : | : | : | : | : | : | : | : | : |  
 OY 263 LNSRVLTVGGES-DSTSLLDKTSIDEKRYVSLRPLPDTGEPLDKTREWQQLFDIMASDP 321  
 ||| : : | : | : | : | : | : | : | : | : | : |  
 Db 176 IDDKIPCMHGGLSPDLSMDQIIRNP-----PVDVPGUICDLSDP 220  
 ||| : : | : | : | : | : | : | : | : | : | : |  
 OY 322 QATMGCVENTLIGAGVNGPDTDNFDFORHRSYVITSHECKNGHFMHDKIIITFSA 381  
 ||| : : | : | : | : | : | : | : | : | : | : | : |  
 Db 221 DKEIDRKGENDRGSVYTFGADVVAEFLQRHDLIDCLRAHQVVEDGYLPFAKROLVLTFS 280  
 ||| : : | : | : | : | : | : | : | : | : | : | : |  
 OY 382 SNNYAIIGSNKAGAYTRINQNLMPHP 405  
 ||| : : | : | : | : | : | : | : | : | : | : |  
 Db 281 PNYCGBFDNAGALMSIDNSIVCSP 304

Search completed: January 20, 2006, 19:48:57  
 Job time : 139 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6  
Copyright (c) 1993 - 2006 Compugen Ltd.

## OM protein - Protein search, using sw model

Run on: January 20, 2006, 19:45:57 ; Search time 46 Seconds

1188.014 Million cell updates/sec

Title: US-09-463-733-1

Perfect score: 3442

Sequence: 1 MDENATRAAITFIQKWRHQQ.....VHDIDPTDCESKVDPKKS 661

Scoring table: BLOSUM62

Gapext 10.0 , Gapext 0.5

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : Issued\_Patents\_AA:\*

- 1: /cgn2\_6/ptodata/1/iaa/5\_COMB\_pep:\*
- 2: /cgn2\_6/ptodata/1/iaa/6\_COMB\_pep:\*
- 3: /cgn2\_6/ptodata/1/iaa/11aa/H\_COMB\_pep:\*
- 4: /cgn2\_6/ptodata/1/iaa/FCITUS\_COMB\_pep:\*
- 5: /cgn2\_6/ptodata/1/iaa/RB\_COMB\_pep:\*
- 6: /cgn2\_6/ptodata/1/iaa/backfiles1.Pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

## #

Result No. Score Query Length DB ID

Description

Sequence 7370, Appl

Sequence 43457, Appl

Sequence 18327, Appl

Sequence 18325, Appl

Sequence 10233, Appl

Sequence 6, Appl

Sequence 873, Appl

Sequence 33, Appl

Sequence 74, Appl

Sequence 7363, Appl

Sequence 1025, Appl

Sequence 6266, Appl

Sequence 8004, Appl

Sequence 6642, Appl

Sequence 12, Appl

Sequence 30, Appl

Sequence 11179, Appl

Sequence 8, Appl

Sequence 276, Appl

Sequence 1018, Appl

Sequence 6, Appl

Sequence 1783, Appl

Sequence 21, Appl

RESULT 1  
US-09-949-016-7370  
; Sequence 7370, Application US/0949016  
; Patient No. 6812339  
; GENERAL INFORMATION:

; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL01307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08

; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FASTSEQ FOR Windows Version 4.0  
; SEQ ID NO 7370  
; LENGTH: 499  
; TYPE: PRT  
; ORGANISM: Human

US-09-949-016-7370

Query Match Best Local Similarity 16.6%; Score 570; DB 2; Length 499;  
Matches 143; Conservative 71; Mismatches 142; Indels 68; Gaps 16;

QY 11 FICKWKRHQRARREMRORCRNWNQFONIERYASSDQASLYKFENDLIMQPAGRKQYQ 70  
95 YIKGYYRRAKSNMVALGK-----FR-----AALRD-----VETVVKYKPHDKDAKRYQ 137

QY 71 -----SSAHR-SVLDDKD---DLVFBG--DIVNAKIELPIRKHIDL 107

QY 138 ECKNIVKQKAFAERIADEHKNSWVDLSDIESTIBYSGPKLDRDKVUTSPMKR--- 193

QY 108 LIDVFRKRCGRNRLAHPKVYALLREAAKSILQKPNISPVASTSQVTVCGDIAKGKDDLL 167

QY 194 LMQWTKOOK--KLHRCQAVQLVQVKSVLSLSTIVTLEKTBTIVCGDTHQGPFDLL 251

QY 1-E8 VVLAHKNGPSSSNPYVINGDFUDRGKQGLLEVLLILSYLAPNPAVTLANGHEDSTDNA 227

QY 252 NIFELNLGPSETMPYINGDFUDRGKQGLLEVLLILSYLAPNPAVTLANGHEDSTDNA 311

QY 228 RICPRIRTESKYPNHRILARIDEVRWLPLGSVLSNRVLTVHGGR--SDSTSUDIKS 285

QY 312 IYGFEGEVAKTY--AQMYLFSEVTFWPLAQCINGKVLMHGGFLSEDGVTLDDIK 368

QY 286 IDRKYVSLRPLTDGBPLDKTEWOOJFDIMWSDPQATMGCVPTLIRGAGVWFGPDVTD 345

RESULT 2  
US-09-517-779-2  
; Sequence 2, Application US/09517779  
; Patent No. 6660511  
; GENERAL INFORMATION:  
; APPLICANT: Xu, Xiang  
; TITLE OF INVENTION: Cell Cycle Proteins Associated with Rad9, Compositions  
; FILE REFERENCE: A6823/RMS/DAV  
; CURRENT APPLICATION NUMBER: US/09/517,779  
; CURRENT FILING DATE: 2000-03-02  
; NUMBER OF SEQ ID NOS: 11  
; SOFTWARE: Patentin Ver. 2.1  
; SEQ ID NO: 2  
; LENGTH: 494  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; US-09-517-779-2

Query Match 15.8%; Score 543; DB 2; Length 494;  
Best Local Similarity 33.4%; Pred. No. 1.1e-43; Gaps 18;  
Matches 142; Conservative 71; Mismatches 142; Indels 70; Gaps 18;

Qy 11 PIQKWMRQRREMQRCWQIQPQNLEYASQDQALYKPFNDLKHMPQAGRNQYQ 70  
Db 90 YIKGYRRRAASNNALGK-----FR-----AALRD-----YETVVKVKPHDKDMKYQO 132  
Qy 71 -----GSAHV-SVLDDKD---DLVZBFG--DIVAKXIBPIRKHIDL 107  
Db 133 EGNKIVKOKAPERAJAGDEHKOSVSDIDESITIEBYSGKLGEBOKVTISPMKE--- 188  
Qy 108 LIDVFKERKERGRNLHAKELKOLPNISPVSTAVSQQTVCGLQJGKUDLL 167  
Db 189 IWKWYDOKK-KLHKCAYOILVQKVKEVLKUSTLVETTLKETEKLITC-DHGPQYDLL 245  
Qy 168 VVUHKNGLPSSPPY-VENGDFDRGKGREGLEUVLILLSLYLAQPAVFLNGRNHEDSVMN 226  
Db 246 NTPELNGLPSPTEWPYDINGDFDVGSPSVVLTGPKLKPDPHILLGNHETDNMN 305  
Qy 227 ARGFIREVESKIPRNHRKILAFIDEVRVWMLPGSVNSRVLVHGGF--SDSTSADLIK 284  
Db 306 QIYGPBGSVKAFT--AQMYELSFSEFWMLPLAQCINGKUIMHGGLPSDGTVDIRR 362  
Qy 285 SIDRGKVSILRPRLTDCEPLDKTEWOOIFPDIMSPDQATMCVCPNTLRGAAGWRFDPDT 344  
Db 363 KLERNR---QPP---DGPMP-----CDLWSDPQOPONG-RSISKRGVSCQGPDV 406  
Qy 345 DNFLQRHLRSYVVRSHCKPNSHPMDNKTITFSASNYYAIGSNKGAYIRL-NNQLMP 403  
Db 407 KAPLEENNLDYIIRSHEVTAEGYEVAHGGRCVTVFASPNCIDOMGKASYTHLQSDLRP 466  
Qy 404 HFVQY 408  
Db 467 QFHQF 471

RESULT 2  
US-09-270-767-43457  
; Sequence 2, Application US/09270767  
; Patent No. 6703491

RESULT 3  
US-09-270-767-43457  
; Sequence 43457, Application US/09270767  
; Patent No. 6703491

RESULT 4  
US-09-248-796A-18327  
; Sequence 18327, Application US/09248796A  
; Patent No. 674137  
; GENERAL INFORMATION:  
; APPLICANT: Keith Weinstock et al  
; TITLE OF INVENTION: NUCLEAR ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN  
; TITLE OF INVENTION: NUCLEIC ACIDS AND PROTEINS OF DROSOPHILA MELANOGASTER  
; FILE REFERENCE: 7326-0094  
; CURRENT APPLICATION NUMBER: US/09/248,796A  
; CURRENT FILING DATE: 1999-03-17  
; PRIORITY APPLICATION NUMBER: US 60/074,725  
; PRIORITY FILING DATE: 1998-02-13  
; PRIORITY APPLICATION NUMBER: US 60/096,409  
; PRIORITY FILING DATE: 1998-08-13  
; NUMBER OF SEQ ID NOS: 28208  
; SEQ ID NO: 18327  
; LENGTH: 445  
; TYPE: PRT  
; ORGANISM: Candida albicans  
; FEATURE:  
; NAME/KEY: UNSURE  
; LOCATION: (220) (239)  
; OTHER INFORMATION: Identity of amino acid sequences at the above locations are unknc



GENERAL INFORMATION:  
; APPLICANT: Giot, Loic  
; APPLICANT: Mansfield, Traci A.  
; TITLE OF INVENTION: Protein-Protein Complexes and Method of Using Same  
; FILE REFERENCE: 15966-542  
; CURRENT FILING DATE: 2000-03-29  
; PRIOR APPLICATION NUMBER: 60/127,352  
; PRIORITY APPLICATION NUMBER: 60/178,965  
; PRIORITY FILING DATE: 2000-02-01  
; NUMBER OF SEQ ID NOS: 1387  
; SOFTWARE: CuraPatSeqFormatter Version 0.9  
; SEQ ID NO: 1093  
; LENGTH: 323  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (0)..(0)  
; OTHER INFORMATION: Polypeptide Accession Number P36873  
; US-09-538-092-1093

Query Match  
Best Local Similarity 32.4%; Pred. No. 1.3e-30;  
Matches 105; Conservative 58; Mismatches 126; Indels 35; Gaps 11;  
Qy 102 KHHIDLILDVFRKKRGRNLHKPVYALL--LREAKSLKOLPNISPVSTAVSQQTVCGD 158  
Db 6 KLNLDLSIGRLRVEQCSR-PKVNQVLTENERGRGKLSRSREFLSPILLERBAPKICGD 64  
Qy 159 LHKLDDLVVLAHKGIPSSSNPYVNGDFVDRGKGKLEVLLLSLYLAFFPNAVFLNRG 218  
Db 65 IHQOYDIDLRLFREYGGPPESN-YLFQGDYVDRGKOSLETCILCLIAKIKYPENFLRG 123  
Qy 219 NHEDSYNARYGFIRESKYPRNHRILAFDRYRMLPIGSVNSRVLIVHGRS-DS 277  
Db 124 NHECASINRITYGPFYDCKRY--NIKLMKTFD-CFNCLPLIAATVDEKIFCCHGJSLSPDL 180  
Qy 278 TSDLKLISIDRKYVSLRPLTDGEPLDKWQOQFDIMSDPQATMGCVPTNLRGAGV 337  
Db 181 QSMQRBRM-----PTDVPDGQLCDLWSPDKDQDGENDRGRSP 225  
Qy 338 WFGDMDNFQHRLSYYVTRSHCKEKGFMHDNKITTSASYNAYAIGSNKGAYIRL 397  
Db 226 TFGABVVAFKLHKHDLICRAHQAQVVEDGEYFPFAKQOLVLTSAAPNCGERDNAGAMSV 285  
Qy 398 NNQMPFEPVQVISAASCTK 416  
Db 286 DETLMCSF-QILKPADK 303

RESULT 9  
US-09-538-092-873  
Sequence 873, Application US/09538092  
; PATENT NO. 6753314  
; GENERAL INFORMATION:  
; APPLICANT: Giot, Loic  
; APPLICANT: Mansfield, Traci A.  
; TITLE OF INVENTION: Protein-Protein Complexes and Method of Using Same  
; FILE REFERENCE: 15966-542  
; CURRENT APPLICATION NUMBER: US/09/538,092  
; PRIOR APPLICATION NUMBER: 60/127,352  
; PRIORITY FILING DATE: 1999-04-01  
; PRIORITY APPLICATION NUMBER: 60/178,965  
; PRIORITY FILING DATE: 2000-02-01  
; NUMBER OF SEQ ID NOS: 1387  
; SOFTWARE: CuraPatSeqFormatter Version 0.9  
; SEQ ID NO: 873  
; LENGTH: 330  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURES:  
; NAME/KEY: misc\_feature  
; LOCATION: (0)..(0)  
; OTHER INFORMATION: Polypeptide Accession Number P08129  
; US-09-538-092-873

Query Match  
Best Local Similarity 31.7%; Pred. No. 2.1e-30;  
Matches 101; Conservative 60; Mismatches 133; Indels 25; Gaps 8;  
Qy 102 KHHIDLILDVFRKKRGRNLHKPVYALL--LREAKSLKOLPNISPVSTAVSQQTVCGD 158  
Db 6 KLNLDLSIGRLRVEQCSR-PKVNQVLTENERGRGKLSRSREFLSPILLERBAPKICGD 64  
Qy 159 LHKLDDLVVLAHKGIPSSSNPYVNGDFVDRGKGKLEVLLLSLYLAFFPNAVFLNRG 218  
Db 65 IHQOYDIDLRLFREYGGPPESN-YLFQGDYVDRGKOSLETCILCLIAKIKYPENFLRG 123  
Qy 219 NHEDSYNARYGFIRESKYPRNHRILAFDRYRMLPIGSVNSRVLIVHGRS-DS 277  
Db 124 NHECASINRITYGPFYDCKRY--NIKLMKTFD-CFNCLPLIAATVDEKIFCCHGJSLSPDL 180  
Qy 278 TSDLKLISIDRKYVSLRPLTDGEPLDKWQOQFDIMSDPQATMGCVPTNLRGAGV 337

Db 181 QSMEQIRRIMR-----PTDVPDGCLCDLWSDPDKDVQGMGRNDRGVSF 225  
 Qy 338 WREGPDVTNPLQLQRHLSTVRSHCKPGRHGFEPHDNKITTSASNYAIGNSKGAVRL 397  
 Db 226 TFGAEVUVAKPLKHDIDLCRAHQVVEDGYEFFAKROLVTLFSAPNYCGBFDNAGAMSV 285  
 Qy 398 NNQLMMPHTVQYISAASTK 416  
 Db 286 DGTLMCSF-QILKPADK 303

RESULT 10  
 US-09-744-016A-33  
 ; Sequence 33, Application US/09744016A  
 ; Patent No. 6875581  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Dr. Voelkel, Helge  
 ; TITLE OF INVENTION: Method for screening of modulators of calcineurin  
 ; TITLE OF INVENTION: activity  
 ; FILE REFERENCE: A24157PCT  
 ; CURRENT APPLICATION NUMBER: US/09/744,016A  
 ; CURRENT FILING DATE: 2001-10-03  
 ; PRIOR APPLICATION NUMBER: EP08113876  
 ; PRIOR FILING DATE: 1998-07-22  
 ; NUMBER OF SEQ ID NOS: 35  
 ; SOFTWARE: PatentIn Ver. 2.1  
 ; SEQ ID NO 33  
 ; LENGTH: 524  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 ; US-09-744-016A-33

Query Match 11.7%; Score 403; DB 2; Length 524;  
 Best Local Similarity 25.3%; Pred. No. 5.1e-30;  
 Matches 136; Conservative 78; Mismatches 202; Indels 122; Gaps 16;

Qy 93 NAKTBLPLTRRNKHIDLLIDVFRKCRGNRLHAKYVALLIREAKASLKLQPNISPVSTAVSQ 152  
 Db 46 NGKPKVYLKHL-----VREG-RABEVALKINGAIIHQ---EKTMIEDAP 92

Qy 153 VTVCGDLGKDDPLLVLAHKNGLPSSSNPYVENGDFDVRGKGLEVLILLSLYLAPNA 212  
 Db 93 ITVCGDLGQFFDLMKUFEVGGSPSNTR-YFLGDYVDRGYSIECYMLWSLKHPT 151

Qy 213 VELNRGHEDSVNARICPIRETESKYPHRNLAFIDEVVRWLPGSVALSRVLIWQ 272  
 Db 152 LFLLRGNHTECHRLTDYFTPKQBCRIK---SBOVYDAGMETFDCLPLAALQMLQPLCVHG 208

Qy 273 GFG-DSFLSLDKLKSIDRGKYSVSLRPPLTGDBPLDKTYEWQQFDIMMSDPOATMG---- 326  
 Db 209 GMSPERITSLDIKRKLDR---FTEPP-----AFGPVCDLWNSDPSDYGNEKTL 253

Qy 327 --CVPNTLIRGAGTWFGRDPVTNPLQLQRHLSTYVIRSHCKPGRHGFEPHDNK----ITI 378  
 Db 254 EHTHTNTVRCGCSYFSPYPAVCEPLQNNNLISITRAHEAQDAGYRMYSQATGPPSLTI 313

Qy 379 PSASNMYAIGNSKGAYIRLNQMLMPHFVQYISAASTKRLSFQRMGIVBSSALKELAVR 438  
 Db 314 FSPAPNLYDYYNNKAVAKVYENNM----- 337

Qy 439 MRDHDRDELEDFERYKDPKDGSYITISHWCKMENVTKLGLPWRLLRDKLAPGDSQKVN 498  
 Db 338 -----NTRQFNCSPHY----WLNPMDPFTWLSPP-----VQVKVTEMLV 374

Qy 499 NRFLDLDLDTDVLKERAEG-MSYMDALYANK---ASLVAIFIIDANSGRITDEFET 553  
 Db 375 NVLNICSDDBLISDDEBAGSTTVERKEITRKIAKGKARVSIROESBSVILKGLTP 434

Qy 554 ATDILVAVMPPGAVSKAEMLKCRMMDINGDGVKDLNEFLRFLSLHJKRQDENIR 611  
 Db 435 TGTLPLGLVLSGGKQTIE-----TAQEEABERERATGSPLOHKIRSPEAR 480

RESULT 11  
 US-09-487-558B-274  
 ; Sequence 274, Application US/09487558B  
 ; Patent No. 6949356  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Busby, Robert  
 ; APPLICANT: Cali, Brian  
 ; APPLICANT: Hecht, Peter  
 ; APPLICANT: Holtzman, Doug  
 ; APPLICANT: Madden, Kevin  
 ; APPLICANT: Maxon, Mary  
 ; APPLICANT: Milne, Todd  
 ; APPLICANT: No. 6849356man, Thea  
 ; APPLICANT: Royer, John  
 ; APPLICANT: Salama, Sofie  
 ; APPLICANT: Sherman, Amir  
 ; APPLICANT: Silva, Jeff  
 ; APPLICANT: Summers, Eric  
 ; TITLE OF INVENTION: Methods for Improving Secondary Metabolite Production in Fungi  
 ; FILE REFERENCE: 109272-130  
 ; CURRENT APPLICATION NUMBER: US/09/487-558B  
 ; CURRENT FILING DATE: 2000-01-19  
 ; PRIOR APPLICATION NUMBER: US 60/487,558  
 ; PRIOR FILING DATE: 1999-10-20  
 ; NUMBER OF SEQ ID NOS: 446  
 ; SOFTWARE: PatentIn version 3.0  
 ; SEQ ID NO 274  
 ; LENGTH: 692  
 ; TYPE: PRT  
 ; ORGANISM: Saccharomyces cerevisiae  
 ; US-09-487-558B-274

Query Match 11.7%; Score 403; DB 2; Length 692;  
 Best Local Similarity 28.5%; Pred. No. 8.1e-30;  
 Matches 120; Conservative 71; Mismatches 166; Indels 64; Gaps 12;

Qy 38 EVASEQOQAELEYKPFNDLKKHMPOAAGCRKNOYQGSAVSVLDDKODIVVERGDIVNA--- 94  
 Db 299 KHLISNDDDIENSQQLSNIHASMENVNNDNN-----NITDSKKOPNEEFNDIMQSSGN 350

Qy 95 -----KELPTIKHN-DLLIDV-FRKRGNLHPK----YVALIREAKASLKLQPNI 142  
 Db 351 KNAPKKRKKPIDETTQKLQDLAGYAAKTKNVCLKNBNELQICIKARKEIFLQPSLIEL 410

Qy 143 SVUSTAVSQVQVTCGDLGKDDPLLVLKNGLPSSSNPYVENGDFDVRGKGLEVLILL 202  
 Db 411 SP-----PVKVGDPHQGQEDLRFKCGFPBPSSN-YFLGDYVDRGKQSLETILL 462

Qy 203 ISLYLAPNNAVFLNRGHEDDSVMNARYGPIREVESKYPHRNLAFIDBEVRLPGSV 262  
 Db 463 FCYKIKPENFFLRLGNHECANVTRYKGYFDBCKRC--NIKIWKTFD-TNTLPLAII 519

Qy 263 LMSRVLVIHGGSFSDSTSLLDKLKSIDRGKYSVSLRPPLTGDBPLDKTYEWQQFDIMMSDPO 322  
 Db 520 VAGKIFCVHGGALS-----PVLMNSDETRHVV-----RPTDVPDPGLNDLWNSDPT 565

Qy 323 ATMGCVNTLIRGAGTWFGRDPVTNPLQLQRHLSTYVIRSHCKPGRHGFEPHDNKITTSAS 382  
 Db 566 DSPNRMWDNDRKRSVSYCNKVAINKPLKFGRDLYCPLVCAHMVVEDGYEFNDISLVTWPSAP 625

Qy 383 NYTAIGNSKGAYIRLNQMLMPHFVQYISAASTKRLSFQRMGIVBSSALKELAVRNRDH 442  
 Db 626 NYCBFEDNWGAIVMSVSBEGLCSP-----ELLDPDSSAALKQVMMKGRQE 669

Qy 443 R 443  
 Db 670 R 670

RESULT 12  
 US-09-949-016-7363  
 ; Sequence 7363, Application US/09949016  
 ; Patent No. 6812339





QY 153 VTVCGDILAKGKDDLLVVLVHKGNGLPSSSNPYVNGDFVDRGKGLEVLILLISLYLFPNA 212 ; Sequence 11179, Application US/093949016  
Db 88 ITVCGDILHGQFDMLKLFSEVGSSPSNTR-YLFGLDYVDRGFSIBCVLYLMSLKHPT 146 ; Patent No. 6812339  
; GENERAL INFORMATION:  
QY 213 VFLNRGHEDSVNARYGPIRETESKYPRNHRILAFIDRVRLPLGSVINSRVLVHG 272 ; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: C1001307  
; CURRENT APPLICATION NUMBER: US/09/949, 016  
; CURRENT FILING DATE: 2000-04-14  
; PRIORITY NUMBER: 60/241, 755  
; PRIORITY NUMBER: 60/237, 768  
; PRIORITY NUMBER: 60/231, 498  
; PRIORITY FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FASTSEQ for Windows Version 4.0  
; SEQ ID NO: 11179  
; LENGTH: 535  
; TYPE: PRT  
; ORGANISM: Human  
; US-09-744-016A-30  
; Sequence 30, Application US/09744016A  
; PATENT NO. 6875581  
; GENERAL INFORMATION:  
; APPLICANT: Dr. Voelkel, Helge  
; TITLE OF INVENTION: Method for screening of modulators of calcineurin  
; FILE REFERENCE: A34157PCT  
; CURRENT APPLICATION NUMBER: US/09/744, 016A  
; PRIORITY NUMBER: EP98113876  
; PRIORITY FILING DATE: 1998-07-22  
; NUMBER OF SEQ ID NOS: 35  
; SOFTWARE: FASTSEQ Ver. 2.1  
; SEQ ID NO: 30  
; LENGTH: 514  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; US-09-744-016A-30  
Query Match Best Local Similarity 11.7%; Score 402; DB 2; Length 535;  
Matches 106; Conservative 31.7%; Pred. No. 6.6e-30; Indels 54; Gaps 11;  
QY 93 NAKIELPIRKHHIDLIDVFRKKRGNLHPKVVALTIREAAKSILKOLPNISPVSTAVSQQ 152 ; Sequence 11179, Application US/093949016  
Db 67 NGKPKDVTLKHL-----VKEG-RIEBAVKINDGAIRO---BKTMRFDAP 113  
QY 153 VTVCGDILAKGKDDLLVVLVHKGNGLPSSSNPYVNGDFVDRGKGLEVLILLISLYLFPNA 212 ; Patent No. 6812339  
Db 114 ITVCGDILHGQFDMLKLFSEVGSSPSNTR-YLFGLDYVDRGFSIBCVLYLMSLKHPT 172  
QY 213 VFLNRGHEDSVNARYGPIRETESKYPRNHRILAFIDRVRLPLGSVINSRVLVHG 272 ; Sequence 11179, Application US/093949016  
Db 173 VFLNRGHEDSVNARYGPIRETESKYPRNHRILAFIDRVRLPLGSVINSRVLVHG 229  
QY 273 GFS-DSTSLLDKSIDRQKYSTLRPPLTDGSPBLDKTEWQOIDFIMSDPDTMIG----- 326 ; Sequence 11179, Application US/093949016  
Db 46 NGKPKDVTLKHL-----VKEG-RIEBAVKINDGAIRO---BKTMRFDAP 92  
QY 153 VTVCGDILAKGKDDLLVVLVHKGNGLPSSSNPYVNGDFVDRGKGLEVLILLISLYLFPNA 212 ; Sequence 11179, Application US/093949016  
Db 93 ITVCGDILHGQFDMLKLFSEVGSSPSNTR-YLFGLDYVDRGFSIBCVLYLMSLKHPT 151  
; RESULT 18  
QY 213 VFLNRGHEDSVNARYGPIRETESKYPRNHRILAFIDRVRLPLGSVINSRVLVHG 272 ; Sequence 11179, Application US/093949016  
Db 152 VFLNRGHEDSVNARYGPIRETESKYPRNHRILAFIDRVRLPLGSVINSRVLVHG 208 ; Patent No. 6815187  
; GENERAL INFORMATION:  
QY 273 GFS-DSTSLLDKSIDRQKYSTLRPPLTDGSPBLDKTEWQOIDFIMSDPDTMIG----- 326 ; APPLICANT: Simons, Michael  
; TITLE OF INVENTION: Stimulation of angiogenesis via  
; syndecan-4 cytoplasmic domain signaling pathway  
; NUMBER OF SEQUENCES: 25  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: David Prashker, Esq.  
; STREET: P.O. Box 5387  
; CITY: Magnolia  
; STATE: Massachusetts  
; COUNTRY: USA  
; ZIP: 01930  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.50 inch, 1.40 Mb storage  
; COMPUTER: Dell PC  
; RESULT 19

Page  
6

RESULT 23  
 US-09-744-016A-6  
 ; Sequence 6, Application US/09744016A  
 ; Patent No. 6875581  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Dr. Voelkel, Helge  
 ; TITLE OF INVENTION: Method for screening of modulators of calcineurin  
 ; TITLE OF INVENTION: activity  
 ; FILE REFERENCE: A3415PCT  
 ; CURRENT APPLICATION NUMBER: US/09/744, 016A  
 ; CURRENT FILING DATE: 2001-10-03  
 ; PRIORITY APPLICATION NUMBER: EP98113876  
 ; PRIORITY FILING DATE: 1998-07-22  
 ; NUMBER OF SEQ ID NOS: 35  
 ; SOFTWARE: PatentIn Ver. 2.1  
 ; LENGTH: 510  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 ; FEATURES:  
 ; NAME/KV: misc\_feature  
 ; LOCATION: (0)..(0)  
 ; OTHER INFORMATION: Polypeptide Accession Number: Q082029  
 ; US-09-538-092-1283  
 Query Match 11.4%; Score 391.5; DB 2; Length 510;  
 Best Local Similarity 31.1%; Pred. No. 6.4e-29;  
 Matches 101; Conservative 58; Mismatches 119; Indels 47; Gaps 10;  
 Query 102 KKHIDILIDIVRKKGRRNLRPKVVALTIREAKSLKOLPNISPVSTAVSQVTVCGDHG 161  
 Db 45 KRPVDIKAHLME--GRLEESVALRITEGASILRQBNKLIDDA---PVTVCDDHG 93  
 Query 162 KUDLILVVLHKGOLPSSSNPPYVNGDFVDRGKGKGLYLILLISLYLAFFNAVLANGHE 221  
 Db 99 QFDLMKLFPEVGSPANTR-YFLGLDVFDRGFSIECVLYWALKLYPKTIFLARGHE 157  
 Query 153 CRHLTEYPTFQBECKKY--SERVYDACHMADCPFLAALMNQQLCIVHGLSPINTL 209  
 Db 281 DLIKSITRGKVSIIRPLDGPBLDKTEWQOIFDIMSDPOATMG-----CVNTLR 333  
 Query 210 DDIRKLDRFK---EPP-----AYGPMCDILWSDPLEDFGNEKTQBFHTNTVR 259  
 Db 334 GAGWMPDPDVNPLQFRHRSYVIRSHHECKNGHFMHDNK----IITPSASNYAI 387  
 Query 260 GCSYFYSYPAVCEFLQRNLNLSTIRAHQAOGACTYMRKSGTGFPSSLITFSAPNYLDV 319  
 Db 388 GSNNKGAIVRLNNQMLM-----PH 404  
 Db 320 YNNKAALVKYENNMNIRQFNCSPH 344

RESULT 24  
 US-09-538-092-1283  
 ; Sequence 1283, Application US/09538092  
 ; Patent No. 6753314  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Glot, Loic  
 ; APPLICANT: Mansfield, Traci A.  
 ; TITLE OF INVENTION: Protein-Protein Complexes and Method of Using Same  
 ; FILE REFERENCE: 15966-542  
 ; CURRENT APPLICATION NUMBER: US/09/538, 092  
 ; CURRENT FILING DATE: 2000-03-29  
 ; PRIOR APPLICATION NUMBER: 60/127, 352  
 ; PRIOR FILING DATE: 1999-04-01  
 ; PRIOR APPLICATION NUMBER: 60/178, 965  
 ; PRIOR FILING DATE: 2000-02-01  
 ; NUMBER OF SEQ ID NOS: 1387  
 ; SOFTWARE: CuratPatSeqFormatter Version 0.9  
 ; SEQ ID NO: 1283  
 ; LENGTH: 521  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 ; FEATURES:  
 ; NAME/KV: misc\_feature  
 ; LOCATION: (0)..(0)  
 ; OTHER INFORMATION: Polypeptide Accession Number: Q082029  
 ; US-09-538-092-1283  
 Query Match 11.4%; Score 391.5; DB 2; Length 510;  
 Best Local Similarity 31.1%; Pred. No. 6.4e-29;  
 Matches 101; Conservative 58; Mismatches 119; Indels 47; Gaps 10;  
 Query 102 KKHIDILIDIVRKKGRRNLRPKVVALTIREAKSLKOLPNISPVSTAVSQVTVCGDHG 161  
 Db 40 KRPVDIKAHLME--GRLEESVALRITEGASILRQBNKLIDDA---PVTVCDDHG 93  
 Query 162 KUDLILVVLHKGOLPSSSNPPYVNGDFVDRGKGKGLYLILLISLYLAFFNAVLANGHE 221  
 Db 94 QFDLMKLFPEVGSPANTR-YFLGLDVFDRGFSIECVLYWALKLYPKTIFLARGHE 152  
 Query 222 DSVMNARYGTFIREVEKSYPRNHKILAFIDEYRMLPIGSVNSRVLVHGSFS-DSTSL 280  
 Db 281 DLIKSITRGKVSIIRPLDGPBLDKTEWQOIFDIMSDPOATMG-----CVNTLR 333  
 Query 210 DDIRKLDRFK---EPP-----AYGPMCDILWSDPLEDFGNEKTQBFHTNTVR 259  
 Db 334 GAGWMPDPDVNPLQFRHRSYVIRSHHECKNGHFMHDNK----IITPSASNYAI 387  
 Query 260 GCSYFYSYPAVCEFLQRNLNLSTIRAHQAOGACTYMRKSGTGFPSSLITFSAPNYLDV 319  
 Db 388 GSNNKGAIVRLNNQMLM-----PH 404  
 Db 320 YNNKAALVKYENNMNIRQFNCSPH 344

RESULT 25  
 US-09-744-016A-21  
 ; Sequence 21, Application US/09744016A  
 ; Patent No. 6875581  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Dr. Voelkel, Helge  
 ; TITLE OF INVENTION: Method for screening of modulators of calcineurin  
 ; TITLE OF INVENTION: activity  
 ; FILE REFERENCE: A3415PCT  
 ; CURRENT APPLICATION NUMBER: US/09/744, 016A  
 ; CURRENT FILING DATE: 2001-10-03  
 ; PRIOR APPLICATION NUMBER: EP98113876  
 ; PRIOR FILING DATE: 1998-07-22  
 ; NUMBER OF SEQ ID NOS: 35  
 ; SOFTWARE: PatentIn Ver. 2.1  
 ; SEQ ID NO: 21



CURRENT FILING DATE: 2000-04-14  
; PRIORITY APPLICATION NUMBER: 60/241,755  
; PRIORITY FILING DATE: 2000-10-20  
; PRIORITY APPLICATION NUMBER: 60/237,768  
; PRIORITY APPLICATION NUMBER: 60/231,498  
; PRIORITY FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO: 6461  
; LENGTH: 309  
; TYPE: PRT  
; ORGANISM: Human  
; US-09-949-016-6461

Query Match 11.4%; Score 391; DB 2; Length 309;  
Best Local Similarity 31.4%; Pred. No. 3.1e-29;  
Matches 92; Conservative 63; Mismatches 110; Indels 28; Gaps 7;

RESULT 30  
US-09-949-016-8780  
; Sequence 8780, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.

; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: C1001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIORITY APPLICATION NUMBER: 60/241,755  
; PRIORITY FILING DATE: 2000-10-20  
; PRIORITY APPLICATION NUMBER: 60/237,768  
; PRIORITY FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO: 8780  
; LENGTH: 341  
; TYPE: PRT  
; ORGANISM: Human  
; US-09-949-016-8780

Query Match 11.2%; Score 386; DB 2; Length 341;  
Best Local Similarity 30.8%; Pred. No. 1.1e-28; Matches 98; Conservative 63; Mismatches 133; Indels 24; Gaps 7;

RESULT 29  
US-09-949-016-1135  
; Sequence 11335, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: C1001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIORITY APPLICATION NUMBER: 60/241,755  
; PRIORITY FILING DATE: 2000-10-20  
; PRIORITY APPLICATION NUMBER: 60/237,768  
; PRIORITY FILING DATE: 2000-10-03  
; PRIORITY APPLICATION NUMBER: 60/231,498  
; PRIORITY FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO: 11335  
; LENGTH: 311  
; TYPE: PRT  
; ORGANISM: Human  
; US-09-949-016-1135

Query Match 11.4%; Score 391; DB 2; Length 311;  
Best Local Similarity 31.4%; Pred. No. 3.1e-29;  
Matches 92; Conservative 63; Mismatches 110; Indels 28; Gaps 7;

RESULT 31  
US-09-949-016-11414  
; Sequence 11414, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: C1001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIORITY APPLICATION NUMBER: 60/241,755  
; PRIORITY FILING DATE: 2000-10-20  
; PRIORITY APPLICATION NUMBER: 60/237,768  
; PRIORITY FILING DATE: 2000-10-03  
; PRIORITY APPLICATION NUMBER: 60/231,498  
; PRIORITY FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO: 11414  
; LENGTH: 311  
; TYPE: PRT  
; ORGANISM: Human  
; US-09-949-016-11414

Sequence 11414, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CLO01307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO: 11414  
; LENGTH: 355  
; TYPE: PRT  
; ORGANISM: Human  
; US-09-949-016-11414

Query Match 11.2%; Score 384; DB 2; Length 355;  
Best Local Similarity 27.7%; Pred. No. 1.9e-28; Mismatches 98; Conservative 70; MisMatches 126; Indels 60; Gaps 9;

QY 64 GRKNOYQGSAAHVSVDK-----DDLVREFGDTIWNAKTELPIRKNHIDLIDVFRKKRG 117  
Db 33 GRSERQPRAGGIMDKRTEKRDQWQNEC----- 66

QY 118 NRHPPKVAILREBAKSILQQLPNSPVSTAVSQVQVTYCGDLHKLDLVLWLAHKNGLPS 177  
Db 67 KQLSBSQVKSLCEKAKEILTKESNVQBVRC---PVTVCGDVGQFDHMLFRIGKSP 122

Qy 178 SSNPYVENGDFDVGKRGKLVLLSXLAPPRAVFTANGRNHSDVANARVIREES 237  
Db 123 DTN-YLFMGDYDVDRGYYSVTIVLALKVRYRERITIRGNHESRQTOVQFYPYBCLR 181

Qy 238 KYPKRHKRILAFIDEVRYMPLGSVLSNSRLVIGGFSDS-TSALDKISIDRGKVSILR 296  
Db 182 KY--GNANTWKKYFPLDYLPIALVGDGTCHLGSSIDTDHIALDQFQVPHG 239

Qy 297 PPLTDGEPLKTEWQOIFDINMSDP--QATMGCVPTLARGAGWFGPDVTFNQFLRHLIS 354  
Db 240 P-----MCIDLWSPDDRGWGQISP--RGAGTGFQDSETENHANGLT 281

Qy 355 YVTRSHHECKENGHEFMHDIKLITFSASHMYAAGSNKAYIRANOLMPHVQ 408  
Db 282 LVSRAHQLVMEGYNGWCHDRANVUTIFAPSAPNYCYRCGNOAIIAMELDDTLKVSFLQP 335

RESULT 32  
US-09-949-016-9733  
; Sequence 9733, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CLO01307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO: 9733  
; LENGTH: 312

Query Match 11.1%; Score 383.5; DB 2; length 312;  
Best Local Similarity 33.4%; Pred. No. 1.7e-28; Mismatches 93; Indels 53; Gaps 10;

QY 123 KVVALIURRAAKSILQUP-----NISPYSTAVSQVQVTYCGDLAKL 163  
Db 15 KVY---EIAFLCKYLPENDLKLRCOVCDLLEBENQVST---PVTVCGDINGQF 65

QY 164 DDLLVVLWLAHKNGLPSNSNPYVENGDFDVGKRGKLVLLSXLAPPRAVFTANGRNHSD 223  
Db 66 YDLCFLFRIGQVDPDN-YIFMDFVDRGYSLSFTFYLLAKWAKWDRITLRGNHRS 124

QY 224 VMARYRGITREVSKYPRNHRKILAFIDEVRYMPLGSVLSNSRLVIGGFSDS-DSTSDL 282  
Db 125 QRIQVYGFYDCEOKY--GNANAWRYCTKVFDMQTLVALIDBOILCWVGGSLSPDILQ 182

QY 283 IKGIDRKGVSIILRPLTDGEPUDKTMQFIDIMSPQ--ATMGCVPTLARGAGWFG 340  
Db 183 IRTERNGBI-----PKGAFCDLUWADPDEVDUTWAS.P--RGAGLFG 224

RESULT 33  
US-08-452-722-7  
; Sequence 7, Application US/08452722  
; Patent No. 562163  
; GENERAL INFORMATION:  
; APPLICANT: Scott, John D.  
; TITLE OF INVENTION: Calcineurin Inhibitory Compounds and  
; NUMBER OF SEQ ID NOS: 7  
; CORRESPONDENCE ADDRESS:  
; ADDRESS: Marshall, O'Toole, Gerstein, Murray & Borun  
; STREET: 233 South Wacker Drive, 6300 Sears Tower  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: United States of America  
; ZIP: 60606  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC Compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; APPLICATION NUMBER: US/08/452,722  
; CURRENT APPLICATION DATA:  
; FILING DATE: 30-MAY-1995  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/344,227  
; FILING DATE: 23-NOV-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Williams Jr., Joseph A.  
; REGISTRATION NUMBER: 38,659  
; REFERENCE/DOCKET NUMBER: 27866/G1000  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 312-474-6300  
; TELEFAX: 312-474-0448  
; INFORMATION FOR SEQ ID NO: 7:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 487 amino acid  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-08-452-722-7

Query Match 11.0%; Score 380; DB 1; Length 487;

Best Local Similarity 30.9%; Pred. No. 7.7e-28; Matches 101; Conservative 57; Mismatches 115; Indels 54; Gaps 11;

QY 100 IRKNHIDLLIVFRKKGRRNLRHKPVVALIREAKSLKQLPNISPVSTAVSQVTVGDL 159  
 Db 44 VLNHHL-----VKRG-RVDEBIALRINNEGAIIRR---EKTWIEAPITVGDI 90  
 QY 160 HKGLDILVVLVHKGNGPSSSNPYVNGDFVDRGKGRLGEVLVLLLSLYLAPPNAFLNRGN 219  
 91 HQQPFMLKLFPSVGSANTR-YLFGLDVYDGYFSTCBLXLMVKLILYPTSLFLRGN 149

QY 220 HEDSVMMARYGTVREVEISKYPRNHKRITALFDEVYRPLPLGSVNLNVRLVTHGGES-DST 278  
 Db 150 HECRHILTYFFPKQCKIKY--SERVTEAMEAFDSLPLAAILNQOFLCVHGGISPERH 206  
 QY 279 SLDLKSIDRGKVSTIRRPLTDGERPLDKTEMQIDIMSDPOATMG-----CVPNT 331  
 Db 207 TLDIIRLDRFK---EPP-----AFGPMDLWNSDPSDFGNBKSQEHRSHNT 251

QY 332 LRGAGWFGFPDVTFQRRHSYVRSHECKNGHEFMHDNK----ITIFSAINY 385  
 Db 252 VRGCSYFTNYPAVCBFLQNNLISIRAHQDAGYMRKSQTGPPSLITISAPNYL 311

QY 386 AIGSNKGAYIRLNQML-----PH 404  
 Db 312 DVTNNCAVLIKENVVMNIROFNCSPH 338

RESULT 34

US-08-404-731A-7

; Sequence 7, Application US/08404731A

; Patent No. 574454  
 ; GENERAL INFORMATION:

; APPLICANT: Robert Owen, et al.  
 ; TITLE OF INVENTION: Calcineurin Inhibitory Compounds and  
 ; TITLE OF INVENTION: Anchoring Protein  
 ; NUMBER OF SEQUENCES: 9  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun  
 ; STREET: 233 South Wacker Drive, 6300 Sears Tower  
 ; CITY: Chicago  
 ; STATE: Illinois  
 ; COUNTRY: United States of America  
 ; ZIP: 60606

; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/404,731A  
 ; FILING DATE:  
 ; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US 08/44,227  
 ; FILING DATE: 23-Nov-1994  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Williams Jr., Joseph A.

; REGISTRATION NUMBER: 38,659  
 ; REFERENCE/DOCKET NUMBER: 27866/G1000  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: 312-474-6300  
 ; TELEFAX: 312-474-0448

; INFORMATION FOR SEQ ID NO: 7:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 487 amino acids  
 ; TYPE: amino acid  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: protein

; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: 312-474-6300  
 ; TELEFAX: 312-474-0448  
 ; INFORMATION FOR SEQ ID NO: 7:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 487 amino acids  
 ; TYPE: amino acid  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: protein

; TOPOLGY: linear  
 ; MOLECULE TYPE: protein

Query Match 11.0%; Score 380; DB 1; Length 487;

Best Local Similarity 30.9%; Pred. No. 7.7e-28;  
 US-08-404-731A-7

Matches 101; Conservative 57; Mismatches 115; Indels 54; Gaps 11;

QY 100 IRKNHIDLLIVFRKKGRRNLRHKPVVALIREAKSLKQLPNISPVSTAVSQVTVGDL 159  
 Db 44 VLNHHL-----VKRG-RVDEBIALRINNEGAIIRR---EKTWIEAPITVGDI 90  
 QY 160 HKGLDILVVLVHKGNGPSSSNPYVNGDFVDRGKGRLGEVLVLLSLYLAPPNAFLNRGN 219  
 91 HQQPFMLKLFPSVGSANTR-YLFGLDVYDGYFSTCBLXLMVKLILYPTSLFLRGN 149

QY 220 HEDSVMMARYGTVREVEISKYPRNHKRITALFDEVYRPLPLGSVNLNVRLVTHGGES-DST 278  
 Db 150 HECRHILTYFFPKQCKIKY--SERVTEAMEAFDSLPLAAILNQOFLCVHGGISPERH 206  
 QY 279 SLDLKSIDRGKVSTIRRPLTDGERPLDKTEMQIDIMSDPOATMG-----CVPNT 331  
 Db 207 TLDIIRLDRFK---EPP-----AFGPMDLWNSDPSDFGNBKSQEHRSHNT 251

QY 332 LRGAGWFGFPDVTFQRRHSYVRSHECKNGHEFMHDNK----ITIFSAINY 385  
 Db 252 VRGCSYFTNYPAVCBFLQNNLISIRAHQDAGYMRKSQTGPPSLITISAPNYL 311

QY 386 AIGSNKGAYIRLNQML-----PH 404  
 Db 312 DVTNNCAVLIKENVVMNIROFNCSPH 338

RESULT 35

US-08-344-227-7

; Sequence 7, Application US/08344227

; Patent No. 5807693  
 ; GENERAL INFORMATION:

; APPLICANT: Scott, John D.  
 ; TITLE OF INVENTION: Calcineurin Inhibitory Compounds and  
 ; TITLE OF INVENTION: Anchoring Protein  
 ; NUMBER OF SEQUENCES: 7  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun  
 ; STREET: 233 South Wacker Drive, 6300 Sears Tower  
 ; CITY: Chicago  
 ; STATE: Illinois  
 ; COUNTRY: United States of America  
 ; ZIP: 60606

; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/344,227  
 ; FILING DATE:  
 ; CLASSIFICATION: 530

; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Williams Jr., Joseph A.

; REGISTRATION NUMBER: 38,659  
 ; REFERENCE/DOCKET NUMBER: 27866/G1000  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: 312-474-6300  
 ; TELEFAX: 312-474-0448

; INFORMATION FOR SEQ ID NO: 7:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 487 amino acids  
 ; TYPE: amino acid  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: protein

; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: 312-474-6300  
 ; TELEFAX: 312-474-0448  
 ; INFORMATION FOR SEQ ID NO: 7:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 487 amino acids  
 ; TYPE: amino acid  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: protein

; TOPOLGY: linear  
 ; MOLECULE TYPE: protein

Query Match 11.0%; Score 380; DB 1; Length 487;  
 Best Local Similarity 30.9%; Pred. No. 7.7e-28;  
 US-08-404-731A-7

Db 44 VLNKHL-----VKEG-RVDEBIALRINNEGAILLR---EKTMEVEAPITVGCDI 90  
 Qy 160 HCKLDDLUVLUHKNGLPPSSNPVPGFPUVDRSKRGLEVLILLSLVIAAPNAFLNRGN 219  
 Db 91 HQGPFDMLKLFEVGGSPANTR-YLFGLQYDVGYSFSTCVCYLWVKLYPSFLRGN 149  
 Qy 220 HEDSVNARYGFIREVESKYPRMKRILAFITDEVRMPLGSVLSNSRLVINGGFS-DST 278  
 Db 150 HEGRHLTYFTFKOECIKY--SERVYACMEAPDSPLAALNNQFLCHVHGLSPRH 206  
 Qy 279 SLDLKSIDRGKVSILRPLPDTGEPLDKTEWQIDPMWSDPQATNG-----CVPT 331  
 Db 207 TLDDIRLDRFK----EPP-----AFGPMCDLWSDSEDEGNEKSQBHFHSNT 251  
 Qy 312 LRGAGVWFGPDVTNPIORHRSYVIRSHECKNGHEFMHDK-----ITIFSAHY 385  
 Db 252 VRGCSYFPNVPAVCEFLQNNLISITRAHEADQAGYMRKSOTTGFPSSLITIFSPNL 311  
 Qy 386 AIGSNKGAVIRLNLQM-----PH 404  
 Db 312 DVVNKAVALKYENNNVNRQFCSPH 338

RESULT 36  
 US-08-503-226B-7  
 ; Sequence 7, Application US/08503226B  
 ; Patent No. 5871945  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Lockerbile, Robert Owen, et al.  
 ; TITLE OF INVENTION: Calcineurin Inhibitory Compounds and  
 ; NUMBER OF SEQUENCES: 42  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun  
 ; STREET: 233 South Wacker Drive, 6300 Sears Tower  
 ; CITY: Chicago  
 ; STATE: Illinois  
 ; COUNTRY: United States of America  
 ; ZIP: 60606  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC Compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/503,226B  
 ; FILING DATE:  
 ; CLASSIFICATION: 436  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US 08/404,731  
 ; FILING DATE: 15-MAR-1995  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US 08/344,227  
 ; FILING DATE: 23-NOV-1994  
 ; ATTORNEY/ AGENT INFORMATION:  
 ; NAME: Williams Jr., Joseph A.  
 ; REGISTRATION NUMBER: 38,659  
 ; REFERENCE/DOCKET NUMBER: 27866/32861  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: 312-474-6300  
 ; TELEFAX: 312-474-0448  
 ; INFORMATION FOR SEQ ID NO: 7:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 487 amino acids  
 ; TYPE: amino acid  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: protein  
 ; US-08-503-226B-7

RESULT 37  
 US-08-721-458B-7  
 ; Sequence 7, Application US/08721458B  
 ; Patent No. 6107104  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Lockerbile, Robert Owen, et al.  
 ; TITLE OF INVENTION: Calcineurin Inhibitory Compounds and  
 ; NUMBER OF SEQUENCES: 74  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun  
 ; STREET: 233 South Wacker Drive, 6300 Sears Tower  
 ; CITY: Chicago  
 ; STATE: Illinois  
 ; COUNTRY: United States of America  
 ; ZIP: 60606  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/721,458B  
 ; FILING DATE:  
 ; CLASSIFICATION: 436  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US 08/404,731  
 ; FILING DATE: 15-MAR-1995  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US 08/344,227  
 ; FILING DATE: 23-NOV-1994  
 ; ATTORNEY/ AGENT INFORMATION:  
 ; NAME: Williams Jr., Joseph A.  
 ; REGISTRATION NUMBER: 38,659  
 ; REFERENCE/DOCKET NUMBER: 27866/3276  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: 312-474-6300  
 ; TELEFAX: 312-474-0448  
 ; INFORMATION FOR SEQ ID NO: 7:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 487 amino acids  
 ; TYPE: amino acid  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: protein  
 ; US-08-721-458B-7

Query Match 11.0%; Score 380; DB 1; Length 487;  
 Best Local Similarity 30.9%; Pred. No. 7, 7e-28;  
 Matches 101; Conservative 57; Mismatches 115; Indels 54; Gaps 11;

Query Match 11.0%; Score 380; DB 2; Length 487;  
 Best Local Similarity 30.9%; Pred. No. 7.7e-28; Mismatches 115; Indels 54; Gaps 11;  
 Matches 101; Conservative 57; Mismatches 115; Indels 54; Gaps 11;

QY 100 IRKHNHIDLIVFRKRGRGNRHPKVVALIRRAAKSLQKOLPNISPVSTAVSQVTWCGDL 159  
 :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||:  
 Db 44 VLNKHL-----VKRG-RVDEBIALRINGAATIRR---EKTMIEVARITVGDI 90  
 :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||:  
 QY 160 HQKLDLILVVLHKNGLPSSSNPYVNGDFVRGKGLEVLILLSLYLAPPNAFLRGN 219  
 :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||:  
 Db 91 HQQFDLMLKPLFVGSPANTR-YLFQDYVDRGYFSIECVLYWVKGISPEIH 149  
 :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||:  
 QY 220 HEDSVNARYGFIRESVKYPRNHKRILAFDDEVVRMLPLGSVNSRLVHGGES-DST 278  
 :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||:  
 Db 150 HECRHLTEYFFPKOECIKY--SERVTEAMEAFDSLPLAAILNOQFLCWHGGISPEIH 206  
 :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||:  
 QY 279 SLDLIKSIDRKYVSIRLRPDTGSPDKTREWQQTIDMWDPOATMG-----CVPNT 331  
 :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||:  
 Db 207 TUDDIRRLDRFK----EPP-----AFGPMCDLLMSDPSRDGFNEKSQEHFSHT 251  
 :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||:  
 QY 332 LRGAGIWFGPDTDNFLQRHLISYVRSHECKPNGBHFMHDNK----ITIFASNY 385  
 :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||:  
 Db 252 VRGCSYFTNYPAVCBLFLQNNLISIRAHQDAGTYMRQSQTGPPSLITIFAPNL 311  
 :|||: :|||: :|||: :|||: :|||: :|||: :|||:  
 QY 386 AIGSNKGAYIRLNNQLM-----PH 404  
 :|||: :|||: :|||: :|||: :|||: :|||:  
 Db 312 DVYNNAKAVALKYENNNMIROFNCSPH 338  
 :|||: :|||: :|||: :|||: :|||:  
  
**RESULT 38**  
 US-09-744-016A-9  
 ; Sequence 9, Application US/09744016A  
 ; Patent No. 6075581  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Dr. Voelkel, Helge  
 ; TITLE OF INVENTION: Method for screening of modulators of calcineurin  
 ; TITLE OF INVENTION: activity  
 ; FILE REFERENCE: A3157PCT  
 ; CURRENT APPLICATION NUMBER: US/09/744, 016A  
 ; CURRENT FILING DATE: 2001-10-03  
 ; PRIORITY APPLICATION NUMBER: EP98113876  
 ; PRIORITY FILING DATE: 1998-07-22  
 ; NUMBER OF SEQ ID NOS: 35  
 ; SOFTWARE: PatentIn Ver. 2.1  
 ; SEQ ID NO: 9  
 ; LENGTH: 528  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 ; US-09-744-016A-9  
 ; US-09-949-016-11683  
 Query Match 11.0%; Score 380; DB 2; Length 530;  
 Best Local Similarity 30.9%; Pred. No. 8.9e-28; Mismatches 115; Indels 54; Gaps 11;  
 Matches 101; Conservative 57; Mismatches 115; Indels 54; Gaps 11;

QY 100 IRKHNHIDLIVFRKRGRGNRHPKVVALIRRAAKSLQKOLPNISPVSTAVSQVTWCGDL 159  
 :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||:  
 Db 60 VLNKHL-----VKRG-RVDEBIALRINGAATIRR---EKTMIEVARITVGDI 106  
 :|||: :|||: :|||: :|||: :|||: :|||: :|||:  
 QY 160 HQKLDLILVVLHKNGLPSSSNPYVNGDFVRGKGLEVLILLSLYLAPPNAFLRGN 219  
 :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||:  
 Db 107 HQQFDLMLKPLFVGSPANTR-YLFQDYVDRGYFSIECVLYWVKGISPEIH 165  
 :|||: :|||: :|||: :|||: :|||: :|||:  
 QY 220 HEDSVNARYGFIRESVKYPRNHKRILAFDDEVVRMLPLGSVNSRLVHGGES-DST 278  
 :|||: :|||: :|||: :|||: :|||: :|||: :|||:  
 Db 166 HECRHLTEYFFPKOECIKY--SERVTEAMEAFDSLPLAAILNOQFLCWHGGISPEIH 222  
 :|||: :|||: :|||: :|||: :|||: :|||:  
 QY 279 SLDLIKSIDRKYVSIRLRPDTGSPDKTREWQQTIDMWDPOATMG-----CVPNT 331  
 :|||: :|||: :|||: :|||: :|||: :|||:  
 Db 223 TUDDIRRLDRFK----EPP-----AFGPMCDLLMSDPSRDGFNEKSQEHFSHT 267  
 :|||: :|||: :|||: :|||: :|||:  
 QY 332 LRGAGIWFGPDTDNFLQRHLISYVRSHECKPNGBHFMHDNK----ITIFASNY 385  
 :|||: :|||: :|||: :|||: :|||:  
 Db 328 DVYNNAKAVALKYENNNMIROFNCSPH 354  
 :|||: :|||: :|||: :|||:  
  
**RESULT 40**  
 US-09-744-016A-24  
 ; Sequence 24, Application US/09744016A  
 ; Patent No. 6075581  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Dr. Voelkel, Helge  
 ; TITLE OF INVENTION: Method for screening of modulators of calcineurin  
 ; FILE REFERENCE: A3157PCT  
 ; CURRENT APPLICATION NUMBER: US/09/744, 016A  
 ; PRIORITY APPLICATION NUMBER: EP98113876

```

; PRIORITY FILING DATE: 1998-07-22
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 24
; LENGTH: 535
; TYPE: PROTEIN
; ORGANISM: Homo sapiens
; US-09-744-016A-24

```

```

Query Match          11.0%; Score 380; DB 2; Length 535;
Best Local Similarity 30.9%; Pred. No. 9.1e-28; Mismatches 115; Indels 54; Gaps 11;
Matches 101; Conservative 57; MisMatches 115; Indels 54; Gaps 11;
Qy 100 IRKHNHIDLIVDVERKKRGRNLHPKVVALIREAKSKIQOLPNISPVSTAVSQVTGGL 159
Db 65 VLRKHL-----VKEG-RVDEBIALRHLNEGATLLR---EKTMLVEAPITVGDI 111
Qy 160 HSGKLDLIVWLHNGLSSSNPVIWNGFDVDRSGKEVLILLSLVIAFPNVAFLRG 219
Db 112 HGQPFPLMKLPEYGGSPANTR-YLFGLSYVDRGYSFECVLYIWLKLYPSLFLRG 170
Qy 220 HEPSVMNARYGRTREVEKYPRNHKRLIAFDIDEVYRMWLPGSVANSRVLIVHGGFS-DST 278
Db 171 HECHLHESYTFKQCKRY--SERVIEACMADPSLPLAHLNQQLCVOHGLSPRIH 227
Db 279 SLDLKLSDRKGVSILRPLPLTGPPLDKTMQIQIPDMWSPQATMG-----CVPNT 331
Db 228 TUDDIRRDRFK-----EPP-----ARGPMCDLWSPDSEDRENKSQBHFSHNT 272
Qy 332 LRGAGWPGDPDVNFQKRLSYVVISHECKNGHESFMNK-----ITIFSAENY 385
Db 273 VRGCSYFVNYPAVECEFLQNNLISIRAHADGAYMRKSQTTGFSPLITTSAPYL 332
Qy 386 AIGSNKGAYIRLNQML-----PH 404
Db 333 DVINKKAVLVKENNMINIRQFNCSPH 359

```

Search completed: January 20, 2006, 19:53:38  
Job time : 48 secs

THIS PAGE BLANK (USPTO)



Db 121 VSTAVSQOVTGCDLHGKLDLIVVLRHGLPSSSNPYVFNEDFVDRGKRLBLVLLLS 180  
 Qy 205 LYLAFFNAPVFLURGNHEDSVWANRGTETRESKYPRHKRLAIDEVYRKLPLGSVIN 264  
 Db 181 LYLAFFNAPVFLURGNHEDSVWANRGTETRESKYPRHKRLAIDEVYRKLPLGSVIN 240  
 Qy 265 SRLVLVHGGPSSTSDLIKSDIRGKVSLIRPLTGDPDQPTEMOQIFDTMWSPOAT 324  
 Db 241 SRLVLVHGGPSSTSDLIKSDIRGKVSLIRPLTGDPDQPTEMOQIFDTMWSPOAT 300  
 Qy 325 MCGVPNTLGRAGWMFGPDVTDIFLQRRLSYVIRSHECKPKNGRFEMDNKLTIFASNY 384  
 Db 301 MCGVPNTLGRAGWMFGPDVTDIFLQRRLSYVIRSHECKPKNGRFEMDNKLTIFASNY 360  
 Qy 385 YAGSGNGAYILNLQNMPLMFQYIQAQSOTKLSQFORMGVESSALKELAVRMDHRD 444  
 Db 361 YAGSGNGAYILNLQNMPLMFQYIQAQSOTKLSQFORMGVESSALKELAVRMDHRD 420  
 Qy 445 LDEDEFRKYDPDSGYISISHCKVMENVTKGPLWRLRDKLAPGTDOSQNKNTLDL 504  
 Db 421 LDEDEFRKYDPDSGYISISHCKVMENVTKGPLWRLRDKLAPGTDOSQNKNTLDL 480  
 Qy 505 LDTDVILAEADGMWSNDALYANKASLVAIFIIDANSGTLDRETAIDLVAHMPG 564  
 Db 481 LDTDVILAEADGMWSNDALYANKASLVAIFIIDANSGTLDRETAIDLVAHMPG 540  
 Qy 565 AYSKAEMLEKCRMMDNGDKDNLTEAFLSDLRKEDENIRRSTGSPAKTA 624  
 Db 541 AYSKAEMLEKCRMMDNGDKDNLTEAFLSDLRKEDENIRRSTGSPAKTA 600  
 Qy 625 TDPVTLADKSKNTLVHEDIDPTCESKVDPKKS 661  
 Db 601 TDPVTLADKSKNTLVHEDIDPTCESKVDPKKS 637

**RESULT 2**  
 US-10-369-493-5651  
 ; Sequence 5651, Application US/10369493  
 ; Publication No. US20030233675A1  
 ; GENERAL INFORMATION:  
 -  
 ; APPLICANT: Cao, Yongwei  
 ; APPLICANT: Hinkle, Gregory J.  
 ; APPLICANT: Slater, Steven C.  
 ; APPLICANT: Chen, Xiaofeng  
 ; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF FILE REFERENCER: 38-10(52052)B  
 ; CURRENT APPLICATION NUMBER: US/10/369, 493  
 ; PRIORITY APPLICATION NUMBER: US 60/360, 039  
 ; PRIORITY FILING DATE: 2002-02-21  
 ; NUMBER OF SEQ ID NOS: 47374  
 ; SEQ ID NO 5651  
 ; LENGTH: 722  
 ; TYPE: PRT  
 ; ORGANISM: Caenorhabditis elegans

US-10-369-493-5651

**RESULT 3**  
 US-10-287-226-324  
 ; Sequence 324, Application US/10287226  
 ; Publication No. US20040086875A1  
 ; GENERAL INFORMATION:  
 -  
 ; APPLICANT: Agee, Michele L.  
 ; APPLICANT: Alsobrook, John P.  
 ; APPLICANT: Bergia, Constance,  
 ; APPLICANT: Boldog, Florence  
 ; APPLICANT: Burgess, Catherine E.,  
 ; APPLICANT: Chant, John S.  
 ; APPLICANT: Chaudhuri, Amitabha,  
 ; APPLICANT: DiPippo, Vincent A.,  
 ; APPLICANT: Edinger, Shlomit R.,  
 ; APPLICANT: Eisen, Andrew,  
 ; APPLICANT: Ellerman, Karen,  
 ; APPLICANT: Gangolli, Esha A.,  
 ; APPLICANT: Gorman, Linda A.,  
 ; APPLICANT: Gerlach, Valerie,  
 ; APPLICANT: Ji, Weizhen,  
 ; APPLICANT: Kekuda, Ramesh,  
 ; APPLICANT: Khramtsov, Nikolai,  
 ; APPLICANT: Li, Li,  
 ; APPLICANT: Malvankar, Uriel M.,  
 ; APPLICANT: MacDougall, John R.,  
 ; APPLICANT: Mezes, Peter S.,  
 ; APPLICANT: Miller, Charles E.,  
 ; APPLICANT: Millet, Isabelle,  
 ; APPLICANT: Ooi, Chean Eng,  
 ; APPLICANT: Ort, Tatjana,  
 ; APPLICANT: Padigaru, Muralidhara,  
 ; APPLICANT: Patturajan, Meera,  
 ; APPLICANT: Rastall, Luca,  
 ; APPLICANT: Rieger, Daniel K.,  
 ; APPLICANT: Rothberg, Mark E.,  
 ; APPLICANT: Shenoy, Suresh G.,  
 ; APPLICANT: Spaderna, Steven K.,

US-10-287-226-324

; Sequence 324, Application US/10287226  
 ; Publication No. US20040086875A1  
 ; GENERAL INFORMATION:  
 -  
 ; APPLICANT: Agee, Michele L.  
 ; APPLICANT: Alsobrook, John P.,  
 ; APPLICANT: Bergia, Constance,  
 ; APPLICANT: Boldog, Florence  
 ; APPLICANT: Burgess, Catherine E.,  
 ; APPLICANT: Chant, John S.  
 ; APPLICANT: Chaudhuri, Amitabha,  
 ; APPLICANT: DiPippo, Vincent A.,  
 ; APPLICANT: Edinger, Shlomit R.,  
 ; APPLICANT: Eisen, Andrew,  
 ; APPLICANT: Ellerman, Karen,  
 ; APPLICANT: Gangolli, Esha A.,  
 ; APPLICANT: Gorman, Linda A.,  
 ; APPLICANT: Gerlach, Valerie,  
 ; APPLICANT: Ji, Weizhen,  
 ; APPLICANT: Kekuda, Ramesh,  
 ; APPLICANT: Khramtsov, Nikolai,  
 ; APPLICANT: Li, Li,  
 ; APPLICANT: Malvankar, Uriel M.,  
 ; APPLICANT: MacDougall, John R.,  
 ; APPLICANT: Mezes, Peter S.,  
 ; APPLICANT: Miller, Charles E.,  
 ; APPLICANT: Millet, Isabelle,  
 ; APPLICANT: Ooi, Chean Eng,  
 ; APPLICANT: Ort, Tatjana,  
 ; APPLICANT: Padigaru, Muralidhara,  
 ; APPLICANT: Patturajan, Meera,  
 ; APPLICANT: Rastall, Luca,  
 ; APPLICANT: Rieger, Daniel K.,  
 ; APPLICANT: Rothberg, Mark E.,  
 ; APPLICANT: Shenoy, Suresh G.,  
 ; APPLICANT: Spaderna, Steven K.,

US-10-287-226-324

; Sequence 324, Application US/10287226  
 ; Publication No. US20040086875A1  
 ; GENERAL INFORMATION:  
 -  
 ; APPLICANT: Agee, Michele L.  
 ; APPLICANT: Alsobrook, John P.,  
 ; APPLICANT: Bergia, Constance,  
 ; APPLICANT: Boldog, Florence  
 ; APPLICANT: Burgess, Catherine E.,  
 ; APPLICANT: Chant, John S.  
 ; APPLICANT: Chaudhuri, Amitabha,  
 ; APPLICANT: DiPippo, Vincent A.,  
 ; APPLICANT: Edinger, Shlomit R.,  
 ; APPLICANT: Eisen, Andrew,  
 ; APPLICANT: Ellerman, Karen,  
 ; APPLICANT: Gangolli, Esha A.,  
 ; APPLICANT: Gorman, Linda A.,  
 ; APPLICANT: Gerlach, Valerie,  
 ; APPLICANT: Ji, Weizhen,  
 ; APPLICANT: Kekuda, Ramesh,  
 ; APPLICANT: Khramtsov, Nikolai,  
 ; APPLICANT: Li, Li,  
 ; APPLICANT: Malvankar, Uriel M.,  
 ; APPLICANT: MacDougall, John R.,  
 ; APPLICANT: Mezes, Peter S.,  
 ; APPLICANT: Miller, Charles E.,  
 ; APPLICANT: Millet, Isabelle,  
 ; APPLICANT: Ooi, Chean Eng,  
 ; APPLICANT: Ort, Tatjana,  
 ; APPLICANT: Padigaru, Muralidhara,  
 ; APPLICANT: Patturajan, Meera,  
 ; APPLICANT: Rastall, Luca,  
 ; APPLICANT: Rieger, Daniel K.,  
 ; APPLICANT: Rothberg, Mark E.,  
 ; APPLICANT: Shenoy, Suresh G.,  
 ; APPLICANT: Spaderna, Steven K.,

US-10-287-226-324

; Sequence 324, Application US/10287226  
 ; Publication No. US20040086875A1  
 ; GENERAL INFORMATION:  
 -  
 ; APPLICANT: Agee, Michele L.  
 ; APPLICANT: Alsobrook, John P.,  
 ; APPLICANT: Bergia, Constance,  
 ; APPLICANT: Boldog, Florence  
 ; APPLICANT: Burgess, Catherine E.,  
 ; APPLICANT: Chant, John S.  
 ; APPLICANT: Chaudhuri, Amitabha,  
 ; APPLICANT: DiPippo, Vincent A.,  
 ; APPLICANT: Edinger, Shlomit R.,  
 ; APPLICANT: Eisen, Andrew,  
 ; APPLICANT: Ellerman, Karen,  
 ; APPLICANT: Gangolli, Esha A.,  
 ; APPLICANT: Gorman, Linda A.,  
 ; APPLICANT: Gerlach, Valerie,  
 ; APPLICANT: Ji, Weizhen,  
 ; APPLICANT: Kekuda, Ramesh,  
 ; APPLICANT: Khramtsov, Nikolai,  
 ; APPLICANT: Li, Li,  
 ; APPLICANT: Malvankar, Uriel M.,  
 ; APPLICANT: MacDougall, John R.,  
 ; APPLICANT: Mezes, Peter S.,  
 ; APPLICANT: Miller, Charles E.,  
 ; APPLICANT: Millet, Isabelle,  
 ; APPLICANT: Ooi, Chean Eng,  
 ; APPLICANT: Ort, Tatjana,  
 ; APPLICANT: Padigaru, Muralidhara,  
 ; APPLICANT: Patturajan, Meera,  
 ; APPLICANT: Rastall, Luca,  
 ; APPLICANT: Rieger, Daniel K.,  
 ; APPLICANT: Rothberg, Mark E.,  
 ; APPLICANT: Shenoy, Suresh G.,  
 ; APPLICANT: Spaderna, Steven K.,

US-10-287-226-324

APPLICANT: Spytek, Kimberley A.,  
 APPLICANT: Taugier, Jr., Raymond J.,  
 APPLICANT: Vernet, Corine A.M.,  
 APPLICANT: Zehusen, Bryan D.,  
 APPLICANT: Zhong, Mei

TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEAR ACIDS ENCODING SAME  
 FILE REFERENCE: 21402-480C  
 CURRENT FILING DATE: 2002-11-04  
 PRIOR APPLICATION NUMBER: 60/334,421  
 PRIOR FILING DATE: 2001-11-30  
 PRIOR APPLICATION NUMBER: 60/354,392  
 PRIOR FILING DATE: 2002-02-04  
 PRIOR APPLICATION NUMBER: 60/360,148  
 PRIOR FILING DATE: 2002-02-27  
 PRIOR APPLICATION NUMBER: 60/364,000  
 PRIOR FILING DATE: 2002-03-13  
 PRIOR APPLICATION NUMBER: 60/404,821  
 PRIOR FILING DATE: 2002-08-20  
 PRIOR APPLICATION NUMBER: 60/334,526  
 PRIOR FILING DATE: 2001-11-30  
 PRIOR APPLICATION NUMBER: 60/354,409  
 PRIOR FILING DATE: 2002-02-04  
 PRIOR APPLICATION NUMBER: 60/364,227  
 PRIOR FILING DATE: 2002-03-13  
 PRIOR APPLICATION NUMBER: 60/334,027  
 PRIOR FILING DATE: 2001-11-28  
 PRIOR APPLICATION NUMBER: 60/331,641  
 PRIOR FILING DATE: 2001-11-20  
 Remaining Prior Application data removed - See File Wrapper or PALM.  
 NUMBER OF SEQ ID NOS: 673  
 SEQ ID NO: 324  
 LENGTH: 613  
 TYPE: PRT  
 ORGANISM: Homo sapiens  
 IS-10-287-226-324

Query Match 33.8%; Score 1164.5; DB 4; Length 613;  
 Best Local Similarity 39.3%; Pred. No. 2,8e-87;  
 Matches 247; Conservative 115; Mismatches 200; Indels 67; Gaps 13;

Y 3 EMAIRALIPIQKMYRHRARRENRQRCCRQWQIFQNLVYASEQDQMLRYKEPFDLTKHMPQA 62  
 : ::|||:||||::|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:  
 14 DTSIRALIQLIWTRGKYLKAROHALTIFOSIEYADEQGQWQLSITPPSFMLENYTHI 73

Y 63 AGRKNOYQGSANHVSVDQKD--DLV--BFGDINA-KELPLTRKHNIDLLIDYFRKRG 117  
 : :|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:  
 74 HKKEBLERLARNQSLSRSBQMDRDRWYVDSIDVPDSYNGPRQLPCTDIDILLEKPEQQ- 132

Y 118 NRLHPKVYALLREAAASIKQLNIPISVASTAVSQVTGCGDQGLKGKDDLLVWKGGLS 177  
 : |||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:  
 133 -ILHARYTLEVLFETKVKLQKOMENPHTQITSPSKERVTCIGPLAKGDLLFLIFTKGGLS 191

Y 178 SSNPYVINGDFVDRGKRCRGLREVLLLISLYLAEPNAVAFNGLRGNHEDSVMARYGFTREVES 237  
 : |||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:  
 192 ERNPYVINGDFVDRGKNSIEBILMILCVSPLVYFNDLHARGNHEDPMONLRGYFTKEILH 251

Y 238 KYPRNHRHLILAPIDEVRWLP-----LGSVLNRSRVLIVGGPSSTSLSL 282  
 : |||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:  
 252 KYKLHGRILQILOTFEFKAWLPTENRNRDGTDSKRNKGTVFNA----HGRIK----- 299

Y 283 IKSIDRKVSVSLRPLTDGEP--LDKTEWQOQIDIMSDPQATMGCPNTLRAVGWR 339  
 : |||:|||:|||:|||:|||:|||:|||:|||:|||:  
 300 -----TNGSPTEHLEHNEWBOLDIWSDPKGKGCPTNCRGCGYF 342

Y 340 GDPVTDMFLQRHLISYVRSHECKPKNSHEFMHDNKITIFSFASNYATGSNKGAYIRLN 399  
 : |||:|||:|||:|||:|||:|||:|||:|||:|||:|||:  
 343 GDPVTMKLINKYQKMLRSHECKPKNSHEFMHDNKITIFSFASNYATGSNKGAYIRLN 402

Y 400 QLMRPHFOYISASQTRLSF--KQRMGIVESSAKELAVERMRDHRDELEDERKDYDK 456  
 : |||:|||:|||:|||:|||:|||:|||:|||:  
 403 GTPTRPRFQY---QVTKATCFCPLRQDVTMENSAKILRERKTSKDSLTRAOQDIER 458

RESULT 4  
 US-10-287-226-326  
 Sequence 326, Application US/10287226  
 Publication No. US20040086875A1  
 GENERAL INFORMATION:  
 APPLICANT: Agree, Michele L.,  
 APPLICANT: Albabrook, John P.,  
 APPLICANT: Berghs, Constance,  
 APPLICANT: Boldog, Ferencie,  
 APPLICANT: Burgess, Catherine E.,  
 APPLICANT: Chant, John S.,  
 APPLICANT: Chaudhuri, Amitabha,  
 APPLICANT: DiPippo, Vincent A.,  
 APPLICANT: Edinger, Shlomit R.,  
 APPLICANT: Elser, Andrew,  
 APPLICANT: Ellerman, Karen,  
 APPLICANT: Gangolli, Baha A.,  
 APPLICANT: Gerlach, Valerie,  
 APPLICANT: Ji, Weizhen,  
 APPLICANT: Kekuda, Ramesh,  
 APPLICANT: Kramtsov, Nikolai,  
 APPLICANT: Li, Li,  
 APPLICANT: Malynkar, Uriel M.,  
 APPLICANT: Macbongali, John R.,  
 APPLICANT: Mezes, Peter S.,  
 APPLICANT: Miller, Charles E.,  
 APPLICANT: Millet, Isabelle,  
 APPLICANT: Ooi, Chean Eng,  
 APPLICANT: Ort, Tatiana,  
 APPLICANT: Padigaru, Muralidhara,  
 APPLICANT: Patturajan, Meera,  
 APPLICANT: Rastelli, Luca,  
 APPLICANT: Rieger, Daniel K.,  
 APPLICANT: Rothenberg, Mark B.,  
 APPLICANT: Shenoy, Suresh G.,  
 APPLICANT: Spaderina, Steven K.,  
 APPLICANT: Spytek, Kimberley A.,  
 APPLICANT: Taugier, Jr., Raymond J.,  
 APPLICANT: Vernet, Corine A.M.,  
 APPLICANT: Zehusen, Bryan D.,  
 APPLICANT: Zhong, Mei

TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEAR ACIDS ENCODING SAME  
 FILE REFERENCE: 21402-480C  
 CURRENT FILING DATE: 2002-02-04  
 PRIOR APPLICATION NUMBER: US/10/287,226  
 CURRENT FILING DATE: 2002-11-04  
 PRIOR APPLICATION NUMBER: 60/334,421  
 PRIOR FILING DATE: 2002-02-27  
 PRIOR APPLICATION NUMBER: 60/360,148  
 PRIOR FILING DATE: 2002-03-13  
 PRIOR FILING DATE: 2002-08-20  
 PRIOR APPLICATION NUMBER: 60/334,526  
 PRIOR FILING DATE: 2001-11-30  
 PRIOR APPLICATION NUMBER: 60/354,409

QY 457 DGYVISHWCKMENTVKGFLPWLKARLTDOKWNNTFLDJD-T-VLRABA 515  
 : |||:|||:|||:|||:|||:|||:  
 Db 459 KSGKLVSQSWMCFEMENTVGLNLFWRSLSNNVNIDONGNVEMSSFQNIRIEKPVQBAHS 518  
 QY 516 DCSMVQALYANKASLVAIFNTIDADNSGRITDEPETAIDLVYAHMPGAVSKAEMLK 575  
 : |||:|||:|||:|||:|||:|||:  
 Db 519 ---TLVETLYRTRSDLBTFNATDTHSGLISVEFRAMWKLFSHHYNVHDDSQVNKA 575  
 QY 576 RMADLNGDGKVDNFELAFRSDLHRKE 604  
 : |||:|||:|||:|||:  
 Db 576 NTMDINKGSDIPNEFLKAFV-VHRYE 602

PRIOR FILING DATE: 2002-02-04  
 PRIOR APPLICATION NUMBER: 60/364,227  
 PRIOR FILING DATE: 2002-03-13  
 PRIOR APPLICATION NUMBER: 60/334,027  
 PRIOR FILING DATE: 2001-11-28  
 PRIOR FILING DATE: 2001-11-20  
 ; REMAINING Prior Application data removed - See File Wrapper or PALM.  
 ; SOFTWARE: CuraSeqList version 0.1  
 ; SEQ ID NO 326  
 ; LENGTH: 613  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 ; US-10-287-226-326

Query Match 33.8%; Score 1164.5; DB 4; Length 613;  
 Best Local Similarity 39.3%; Pred. No. 2.8e-87;  
 Matches 247; Conservative 115; Mismatches 200; Indels 67; Gaps 13;  
 QY 3 ENVIRAAIFIQKTYRHRQARREMRRENWQFONLEYASRQDQABLYKFENDLIKMPQA 62  
 Db 14 DISIRAKAIIQWIRGYKARLAKAROHTALTQSISBYADBCOCOMLSTFESPFMLNTHI 73  
 QY 63 AGRKQYQGSATVSLDDKD--DLVE--ERGDIVNA-KIEBPKRNHIDJLIDVVKRG 117  
 Db 74 HKEBLERLNQSLSEODMRDRMDYVDSDVPSYNGRLPPLCTDIDLLEBAFQO- 132  
 QY 118 NRLHPKVVALIREEAKSILQPNISPVASTVSQVTCGGDAGKDDLLVLHKNGLP 177  
 Db 133 -LHAHTVLEVLPETKTKLKMNPHTIQTSKPSKEVIGDAGKDDLLPFLIPYKQPS 191  
 QY 178 SSNPYVTFNGDFVDRGKGLEVILLISLYLAFFPNAFLNRGHEDSYNAMYGFREVES 237  
 Db 192 ERNPYVFNGDFVDRGKNSIBELMILCVSFLVTPNDLHLRNKHEDPMNLNKGFTKIH 251  
 QY 238 KYPNHRHILAFDEVTRWL-----LGSVLNSRVLVLHGFSDSL 282  
 Db 252 KYLKHGRKILQIBERYAWLPTETNRDHGTDSKHNVGUTVNA----HGRIK----- 299  
 QY 283 IKSIDRGKVVSLRPLTDGER--LDKTEWQOIDFDIMSDPQATMCVCPVILRGAGWF 339  
 - Db 300 .....TNGSPTEHILTEHEWQOIDLWSPRKGNCPPNTRGGCYP 342  
 QY 340 GDPVTDNFLQRHLSYVTRSHCKPNCHEFMHDNKITIFSSNYAIGSGNGAYRLNN 399  
 Db 343 GDPVTSKILNKQVQLMLRSHCKPCEGYICHDGKVVTIFASNYAEGSGNGAYIKCS 402  
 QY 400 OLMPHYOYIASQOTKLSF--KQMGIVVSELLAELAVEMDRDRDDEPERKDPK 456  
 Db 403 GTPPRFROY---QVTKATCFOPLRQVDMENSAKILRKERTVISRSKDSLTAFOQDR 458  
 QY 457 DSOYISSHMKUMENTIKGLPWRLLDKLAPGTDQSOKVNNTLDDLD-T-VILAEA 515  
 Db 459 KSKKLSSQWAFMENTILGNIUPWSISLNLNUIDQGNVETMSFONTRLEKPVQRAHS 518  
 QY 516 DGMVMDALYANKASLVAIFNTIDADNSGETILEDEFESTAIDLVIAHMPGAVSKAEMLEKC 575  
 Db 519 --TLVETLYRYSRSLBISINAIDTDHSGLISVEFRAMKULFSSHNVHLDSDQNLK 575  
 QY 576 RMDLNGSGKVDIANTEFRAFRASDHLRKE 604  
 Db 576 NIMDLINKOGSIDPNEFLAKYY--VHRYE 602

RESULT 5  
 US-10-425-115-347492  
 ; Sequence 347492, Application US/10425115  
 ; Publication No. US20040214272A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: La Rosa, Thomas J.  
 ; APPLICANT: Kovacic, David K.  
 ; APPLICANT: Screen, Steven E  
 ; APPLICANT: Taback, Jack E  
 ; APPLICANT: Cao, Yongwei  
 ; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
 ; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
 ; FILE REFERENCE: 38-21-(53313)<sup>B</sup>  
 ; CURRENT APPLICATION NUMBER: US/10/425,114  
 ; CURRENT FILING DATE: 2003-04-28  
 ; NUMBER OF SEQ ID NOS: 73128  
 ; SEQ ID NO 61206  
 ; LENGTH: 488  
 ; TYPE: PRT  
 ; ORGANISM: Zea mays  
 ; OTHER INFORMATION: Clone ID: LIB3732-029-F9\_FLI.pep  
 ; US-10-425-114-61206

Query Match 16.9%; Score 581; DB 4; Length 488;  
 Best Local Similarity 41.4%; Pred. No. 4.5e-39;  
 Matches 127; Conservative 47; Mismatches 109; Indels 24; Gaps 8;

```

RESULT 7
US-11-097-143-18312
; Sequence 18312, Application US/11097143
; Publication No. US2005020558A1
; GENERAL INFORMATION:
; APPLICANT: Venier, J. Craig
; APPLICANT: et al.
; TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID
; TITLE OF INVENTION: ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE
; TITLE OF INVENTION: DROSOPHILA GENES.
; FILE REFERENCE: CL000728
; CURRENT APPLICATION NUMBER: US/11/097,143
; CURRENT FILING DATE: 2005-04-04
; PRIOR APPLICATION NUMBER: 60/157,832
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: 60/160,191
; PRIOR FILING DATE: 1999-10-19
; PRIOR APPLICATION NUMBER: 60/161,932
; PRIOR FILING DATE: 1999-10-28
; PRIOR APPLICATION NUMBER: 60/164,769
; PRIOR FILING DATE: 1999-11-12
; PRIOR APPLICATION NUMBER: 60/173,363
; PRIOR FILING DATE: 1999-12-28
; PRIOR APPLICATION NUMBER: 60/175,693
; PRIOR FILING DATE: 2000-01-12
; PRIOR APPLICATION NUMBER: 60/184,831
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: 60/191,637
; PRIOR FILING DATE: 2000-03-23
; NUMBER OF SEQ ID NOS: 43108
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 18312
; LENGTH: 520
; TYPE: PRT
; ORGANISM: DROSOPHILA
; US-11-097-143-18312

Query Match 16.8%; Score 577;, DB 6; Length 520;
Best Local Similarity 32.1%; Pred. No. 1.1e-38;
Matches 136; Conservative 83; Mismatches 153; Indels 52; Gaps 15;

QY      4 NAIKA-AITQKWRHRQARREM---QRCRNWQIWFQNLRYASQDQABLYKEPFNDLKK- 57
Db      108 SAVKADPAVLKGYYRRAAHMSIGKFQKQALCDFPVAKCR-PNDKDAKLFTECNKIVEN 166
QY      58 -----HMPQAGRKNOYQGSAHVSUDD-KDLVSEFGDINNAKILBLPIRKNHDL 107
Db      167 RAFFERAIAVDKPEKT-LSEMYSDMNEITBDDYKQQLD-----GKTYKEMKE---- 215

RESULT 7
US-11-097-143-18312
; Sequence 18312, Application US/11097143
; Publication No. US2005020558A1
; GENERAL INFORMATION:
; APPLICANT: Venier, J. Craig
; APPLICANT: et al.
; TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID
; TITLE OF INVENTION: ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE
; TITLE OF INVENTION: DROSOPHILA GENES.
; FILE REFERENCE: CL000728
; CURRENT APPLICATION NUMBER: US/11/097,143
; CURRENT FILING DATE: 2005-04-04
; PRIOR APPLICATION NUMBER: 60/157,832
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: 60/160,191
; PRIOR FILING DATE: 1999-10-19
; PRIOR APPLICATION NUMBER: 60/161,932
; PRIOR FILING DATE: 1999-10-28
; PRIOR APPLICATION NUMBER: 60/164,769
; PRIOR FILING DATE: 1999-11-12
; PRIOR APPLICATION NUMBER: 60/173,363
; PRIOR FILING DATE: 1999-12-28
; PRIOR APPLICATION NUMBER: 60/175,693
; PRIOR FILING DATE: 2000-01-12
; PRIOR APPLICATION NUMBER: 60/184,831
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: 60/191,637
; PRIOR FILING DATE: 2000-03-23
; NUMBER OF SEQ ID NOS: 43108
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 18312
; LENGTH: 520
; TYPE: PRT
; ORGANISM: DROSOPHILA
; US-11-097-143-18312

Query Match 16.8%; Score 577;, DB 6; Length 520;
Best Local Similarity 32.1%; Pred. No. 1.1e-38;
Matches 136; Conservative 83; Mismatches 153; Indels 52; Gaps 15;

QY      4 NAIKA-AITQKWRHRQARREM---QRCRNWQIWFQNLRYASQDQABLYKEPFNDLKK- 57
Db      108 SAVKADPAVLKGYYRRAAHMSIGKFQKQALCDFPVAKCR-PNDKDAKLFTECNKIVEN 166
QY      58 -----HMPQAGRKNOYQGSAHVSUDD-KDLVSEFGDINNAKILBLPIRKNHDL 107
Db      167 RAFFERAIAVDKPEKT-LSEMYSDMNEITBDDYKQQLD-----GKTYKEMKE---- 215

RESULT 8
US-10-450-763-40348
; Sequence 40348, Application US/10450763
; Publication No. US20050196754A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 790CIP3-US
; CURRENT APPLICATION NUMBER: US/10/450,763
; CURRENT FILING DATE: 2003-06-11
; PRIOR APPLICATION NUMBER: PCT/US01/08631
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/649,167
; PRIOR FILING DATE: 2000-08-23
; NUMBER OF SEQ ID NOS: 60736
; SOFTWARE: Custom
; SEQ ID NO 40348
; LENGTH: 500
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE: DOMAIN
; NAME/KEY: DOMAIN
; LOCATION: (405)..(460)
; OTHER INFORMATION: Serine/threonine specific protein phosphatases proteins
; OTHER INFORMATION: domain identified by eMATRIX, accession number B100125D, p-vad
; OTHER INFORMATION: 7.51e-37, raw score of 33.11
; FEATURE: DOMAIN
; NAME/KEY: DOMAIN
; LOCATION: (207)..(468)
; OTHER INFORMATION: Ser/Thr protein phosphatase domain identified by Pfam, accession name SThosphatase, E-value=1.4e-106, PParam score of 100
; OTHER INFORMATION: Xaa = X or * as defined in Table 2
; OTHER INFORMATION: (1)..(500)
; OTHER INFORMATION: Xaa = X or * as defined in Table 2
; OTHER INFORMATION: US-10-450-763-40348

Query Match 16.6%; Score 571;, DB 5; Length 500;
Best Local Similarity 34.8%; Pred. No. 3.1e-38;
Matches 143; Conservative 73; Mismatches 153; Indels 42; Gaps 16;

QY      11 PIQKWMRHRQARREM---RCRNWQIWFQNLRYASQDQABLYKEPFNDLKKMPOAGRK- 67
Db      99 YMGYYRRRAAHMSIGKFQKQALDFYETVKVKPHDKDAKMKYTOECNKVK--PKAFER-- 154

```

QY 68 QYQGSAHV-SVLDKDD---DLVERPG--DIVNAKELPIRKONHDLIDVFRKGGRJL 120  
Db 155 AIADEBHKRSVVPSLDISMTIDESYSPKLEDGKVTSFME---LMQWYDQK---KL 208  
Db 121 HPKVVALIREAKSKLQPNISPVSTAVSQCVTVCGLDAGKUDLVLVKHNGRPLSSN 180  
Db 209 HRKCAQIQIVQKVELSKLSTLVTETLKETEKITVCGTHQGYDILUNIFELNGLPSN 268  
QY 181 PYVNGDFVDRGKRGKLGLEVLLSLYLAFPNAVFLNRGHEDSUMMARYFIREVESKYP 240  
Db 269 PYLFGDFVDRGSFSVEVILTRGPKLXPDHFHLRGKHETMNMQTYGPFGEVKAYT 328  
QY 241 RHHGELATIDEVVRWPLGSVLSNSRVLTHGEP--SSTSLOLKSIDRKVVSILRPP 298  
Db 329 ---ACMYBLFSESVFPEWPLAQCINGKVLIMHGJGFSBRGDVTLDIRKIERNR---OPP 380  
QY 299 LTDGEPBLDKTEWQOQFIDIMSDPOTMGCVNTLRAGAGWFGDVTNDLFLQRILSYTR 358  
Db 381 --DSGPM-----CDLWSDPQFQNG-RSISRGVTCQFGDVTKAFLEENNDVIR 429  
QY 359 SHBCKPNHSFHMHDNKKITIFSAASNYYAIGSNKGAYIRU-NNOLMPHFOY 408  
Db 430 SHEVVAEKGVEAHGRCVTVFSAPNYCQDMGNGKASYIHL-QGSDLPQPHQF 480

RESULT 9  
US-10-437-963-195520  
; Sequence 195520, Application US/10437963  
; Publication No. US20040123343A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovacic, David K.  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Wu, Wei  
; APPLICANT: Boukharov, Andrej A.  
; APPLICANT: Barbazuk, Brad  
; APPLICANT: Li, Ping  
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With Best Local Similarity 36.1%; Pred. No. 3 3e-8;  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement Matches 140; Conservative 55; Mismatches 124; Indels 69; Gaps 12;  
; FILE REFERENCE: 38-21(153221)B  
; CURRENT APPLICATION NUMBER: US/10/437,963  
; CURRENT FILING DATE: 2003-05-14  
; NUMBER OF SEQ ID NOS: 204966  
; SEQ ID NO: 195520  
; LENGTH: 483  
; TYPE: PRT  
; ORGANISM: *Oryza sativa*  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT4530\_91460C.1.pep  
; US-10-437-963-195520

Query Match 16.6%; Score 570.5; DB 5; Length 485;  
Best Local Similarity 34.8%; Pred. No. 3 3e-8;  
Matches 145; Conservative 69; Mismatches 154; Indels 49; Gaps 12;  
Matches 145; Conservative 69; Mismatches 154; Indels 49; Gaps 12;

QY 9 AIRICKWYRHRQHAKRERMR--RCNWQFOMLYBASQDQABLYKPFNDLKHMQAAGRNQVGSARHVSVLDBKDD 83  
QY 77 ARYSKGGYRKGAAATGAMGKPKREALKDFOQKRISNPD-----DATRKLUKECEKAVO 128  
QY 68 QYQGSAHVSLDDKDLVBEFGDVINAKEL---PIRKHDLILIVPRKGK 117  
Db 129 KIRFEATISVGDESRVSKASIDSYTRIISTEPOYTGPRVYGDITLTLDVTKAMLFREPKQK- 187  
QY 118 NRJHRYKVALILREAKSKLQPNISPVSTAVSQCVTVCGLDAGKUDLVLVKHNGRPLSSN 177  
Db 188 -CINHRYAIVQVLOTVQQLRSVPSLVDVAVPDGSHFTVCGDVKHQQYFQDNLNFKLNGLPS 246  
QY 178 SSNPYVNGDFVDRGKRGKLGLEVLLSLYLAFPNAVFLNRGHEDSUMMARYFIREVES 237  
Db 247 ENPYLFGDFVDRGSFSVEVILTRGPKLXPDHFHLRGKHETMNMQTYGPFGEVKAYT 306

RESULT 10  
US-10-732-932-9610  
; Sequence 9610, Application US/10732923  
; Publication No. US20050108791A1  
; GENERAL INFORMATION:  
; APPLICANT: Robertson, Michael D  
; TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES  
; FILE REFERENCE: 38-15(22796)C  
; CURRENT APPLICATION NUMBER: US/10/732,923  
; CURRENT FILING DATE: 2003-12-10  
; PRIOR APPLICATION NUMBER: 10/310,154  
; PRIOR FILING DATE: 2002-12-04  
; NUMBER OF SEQ ID NOS: 24149  
; SEQ ID NO: 9610  
; LENGTH: 485  
; TYPE: PRT  
; ORGANISM: *Lycopersicon esculentum*  
; US-10-732-923-9610

Query Match 16.6%; Score 570.5; DB 5; Length 485;  
Best Local Similarity 36.1%; Pred. No. 3 3e-8;  
Matches 140; Conservative 55; Mismatches 124; Indels 69; Gaps 12;  
Matches 140; Conservative 55; Mismatches 124; Indels 69; Gaps 12;

QY 24 EMORRCWQIIFNTLYASEQDQABLYKPFNDLKHMQAAGRNQVGSARHVSVLDBKDD 83  
Db 142 ESOR---SVABSIDYR-----VEVEPOYAGRIR----- 169  
Db 170 -----GDTV-----DFKSMUDFDQK---NHHKRYAQIVQIYQVLTREMALISLV 214  
QY 144 PYSTAVSQCVTVCGLDAGKUDLVLVKHNGRPLSSNBYVFGDFVDRGKRLSIVLILL 203  
Db 215 DIVVPGKGHTVQGDVHQFYDILNFELNGLPSDNPYFLFGDFVDRGSFSLSIVLTF 274  
QY 204 SLYLAFFPAVAPLNRGHEDSUMMARYFIREVESKVRPNHKQILAFIDEVVRWPLGSV 263  
Db 275 AFKCMCPBIAHJLARGHNSKSHKMKTYGEGEVSKSBIFVLFPA---EVFCCLPLAVT 331  
QY 264 NSRLVHGG-FS-DTSLDLTKSIDRKYVSVLRPPLTDGSPLDKTEWQOQFIDIMSDP 321  
Db 332 NEKVPUVHGGLGSVGDVGLSDRAID---FCEPP-----REGMCLKLNSDP 376  
QY 322 QATMGCVNTLRAGAGWFGDPTDNFQQRHRASVYTASHECKPENGHFMHDNKKITFSA 381  
Db 377 QPQGRGRPSK-ROVGLSFGGDFTKRFQENNDLVLVRSHEVDEGYIEHDOKLITVFS 435  
QY 382 SNTYAIGSNKGAVTRLN-NNOLMPHFOY 408  
Db 436 PNCCDMGNGKAFIRFEADMKPNTF 463

RESULT 11  
US-10-043-487-295  
; Sequence 295, Application US/10043487  
; Publication No. US2003005220A1  
; GENERAL INFORMATION:  
; APPLICANT: HERIGENICS  
; TITLE OF INVENTION: Protein-protein interactions between *Shigella flexneri* polypeptides  
; TITLE OF INVENTION: mammalian polypeptides

FILE REFERENCE: B4778A  
; CURRENT APPLICATION NUMBER: US/10/043,487  
; CURRENT FILING DATE: 2002-04-30  
; PRIORITY APPLICATION NUMBER: US 60/261,130  
; PRIORITY FILING DATE: 2001-01-12  
; NUMBER OF SEQ ID NOS: 561  
; SEQ ID NO: 296  
; LENGTH: 494  
; TYPE: PRT  
; ORGANISM: *Shigella flexneri*  
; US-10-043-487-296

Query Match 16.6%; Score 570; DB 4; Length 494;  
Best Local Similarity 33.7%; Pred. No. 3.7e-38; Matches 143; Conservative 71; Mismatches 142; Indels 68; Gaps 16;

QY 11 FLQKWRKRQARREMRRCMWTQIQLTAESEQDQAKLYKFENDLICKPQAGRNKQ 70  
Db 90 YIKGYYRRASNAWMALGK----FR----AALRD----VETWVKYKPHDKDAKOMCQ 132

QY 71 -----GSARV-SVLDKKD---DLVEERG--DIVNAKIELPKRNHIDL 107  
Db 133 ECKNIVKQKPERAIGDHEKRSVSDSLLISMTEDYSGPKELEGKVTSFMEK--- 188

QY 108 LIDYFRKKRKGRLIPKQVAILRRAKSPLKQLPMSPVSTAVSQQTWCGLDGKDDLL 167  
Db 189 LMQWYKDKQK--KLERKCAYQVLQVKEVLSKLSTLVETTLKETKITVCGDTHQFYDIL 246

QY 168 WUHKNGLASSSNYVYFNGDFVDRGSPSVTEVILTFGFKLVDHPHURGNHETDNNQ 206  
Db 247 NIFBLNLGSLPSBTNYIENGDFVDRGSPSVTEVILTFGFKLVDHPHURGNHETDNNQ 227

QY 228 RYGRTREVESKYPRMKRKLAFIDEVYRMLPLASVNSRVLVHGGF--SDSTSLLDKS 285  
Db 307 IYGEGEVKAKEYT--ACMPLRSEVPEFLPLAQCINSKVLMGGFSEBDVLTDIRK 363

QY 286 IDRKYVSIRRPLDGEDPLKDNWQFIDIMSDPQATMGCTENTLGGWVFGPDVTD 345  
Db 364 IERNR----QPP--DSGPM----CDLWNSDPQFONG-RSISKAVGSCQGPDVTK 407

QY 346 NEQCHRISYVIRHECKNGHFMHDWKLITFSASNYAISNSNGAYIRL ANQMPH 404  
Db 408 AFLBENNLDYIIRSHBVKAGYEVAGHGCCVTFSAPNCQDMQNKASYIHLQGSDLRQ 467

QY 405 FVQY 408  
Db 468 FHQP 471

RESULT 12  
US-10-32-923-9609  
; Sequence 9609, Application US/10732923  
; Publication No. US20050108791A1  
; GENERAL INFORMATION:  
; APPLICANT: Edgerton, Michael D  
; TITLE OF INVENTION: NUCLEOTIDE SEQUENCES CODING SIGNAL TRANSDUCTION  
; TITLE OF INVENTION: COMPONENTS IN DURABLE AND BROAD RANGE RESISTANCE  
; TITLE OF INVENTION: STRATEGIES BASED ON PLANT DEFENCE  
; FILE REFERENCE: 2121-0167P  
; CURRENT APPLICATION NUMBER: US/10/298,638  
; CURRENT FILING DATE: 2002-11-19  
; PRIORITY APPLICATION NUMBER: 10/298,638  
; PRIORITY FILING DATE: 2002-11-19  
; NUMBER OF SEQ ID NOS: 28  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO: 27  
; LENGTH: 482  
; TYPE: PRT  
; ORGANISM: *Lycopersicon esculentum*  
; US-10-298-638-27

Query Match 16.3%; Score 562.5; DB 4; Length 482;  
Best Local Similarity 35.8%; Pred. No. 1.5e-37; Matches 139; Conservative 55; Mismatches 125; Indels 69; Gaps 12;

QY 24 EMORRCMWTQIQLTAESEQDQAKLYKFENDLICKPQAGRNKQYQCSARVSVLDKDD 83  
Db 139 ESRQ----SVADSIDRS-----VEVEPQYAGRKE----- 166

QY 84 LYERFGTIVNAKIELPKRNHIDLIVYRKRKGRLIPKQVAILRRAKSPLKQLNIS 143  
Db 167 ----GIVVTL-----DEVKMLDIDKOK--NLAHKRAYOIVQVLTREMIALPSL 211

QY 144 PISTAVSQVTWGDALKDLLVWAKCQSSSPSYVFKDGFVDRGKQLEVLILL 203  
Db 212 DIVVNBKSGHTVCGDVGQPYDINTFELNGPSDNPYLENGDFVDRGSPSLEVITLP 271

QY 204 SLYLAFRPNAVFLNRGNHDSVMARYCPIREVESKYPRMKRKLAPTDDEVYRMLPLGSV 263  
Db 272 ARKMCMSAHLARGNHESKMKYFEGEVRSKSBSIVLFA--EVFCCLPLAHVINSKVP 328

QY 264 NSRQLVITHG--FS-DSTSLLDKSIDSQKYSTLPLDGTBPLDKTWWOQFDIMSDP 321  
Db 329 NEKVYVVRGLPSVDGKUDSRAIDR----FCEPP-----EGLMCELLNSDQPQPG 373

QY 322 QATMGCTPNTLGGWVFGPDVTDNFORHRSYVRSHECKNGHFMHNKLIITFSA 381

Query Match 16.4%; Score 565; DB 5; Length 556;  
Best Local Similarity 40.2%; Pred. No. 1.2e-37; Matches 130; Conservative 47; Mismatches 114; Indels 32; Gaps 9;

RESULT 14  
US-10-298-638-10  
; Sequence 10, Application US/10298638  
; Publication No. US20030177527A1  
; GENERAL INFORMATION:  
; APPLICANT: HARING, Michel A. et al.  
; TITLE OF INVENTION: NUCLEOTIDE SEQUENCES CODING SIGNAL TRANSDUCTION  
; TITLE OF INVENTION: COMPONENTS IN DURABLE AND BROAD-RANGE RESISTANCE  
; FILE REFERENCE: 2121-0167P  
; CURRENT APPLICATION NUMBER: US/10/298 638  
; CURRENT FILING DATE: 2002-11-19  
; PRIORITY APPLICATION NUMBER: 10/298, 638  
; PRIORITY FILING DATE: 2002-11-19  
; NUMBER OF SEQ ID NOS: 28  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO: 10  
; LENGTH: 485  
; TYPE: PRT  
; ORGANISM: Lycopersicon esculentum  
; US-10-298-638-10

Query Match 16.3%; Score 562.5; DB 4; Length 485;  
Best Local Similarity 35.8%; Pred. No. 1.5e-7; Mismatches 12; Gaps 10;  
Matches 139; Conservative 55; MisMatches 125; Indels 69; Gaps 12;

Qy 24 SHORRCWNOIIFORTLEYASBQDAGABLYKPNFLIKHMPQAARKRKNQYQGSANVSVLDDK 83  
Db 142 BSQR--SVADSDYNS-- 169

Qy 84 LVEERGDPVNAKELPIRKONHIDLLIDYPRKRGGRNHLPKVVALURBAAMSLKOPNIS 143  
Db 170 -----GDVNTL-----DEVTKMLUDPDKOK--NLHKRAYOIVLQTRMLRALPSLV 214

Qy 144 PVSTAVSQVTWCGDLHKGKDDLUVLTAKNGJPLSSSNPYVNGDFVDRGRKGLLEVLL 203  
Db 215 DIVPEKGKHITVGDVHQVQFVOLINTFELNGLSPSENPYVNGDFVDRGSISLEVLLP 274

Qy 204 SYLAFFNAVFLNRGNHEDSVNARYGPIREYESKYPRNHRIALFDEVVRWLPGSVL 263  
Db 275 AKKCMCPSAHLHARGNHRSKMKVGFEGFRSKRKFVLFIA---EVVCCPLAHVI 331

Qy 264 NSRVLVHGG FS-DSTSLDLIIKSIDRQKYSILRPLPDTSEPLDKTEWQOQIDMWSDP 321  
Db 332 NEKVUVHTRGLFSVVDGTVKLSDIRADR----FCBPP-----ERGMCLRLWSDP 376

Qy 322 QATMGCVPTNLIGAGWVFGPDVTDNFTLQRHRLSYVYRSHECKPNGHFRMHNDKITTIFSA 381  
Db 377 QRPGPSPSK RGVLSFGGDVKRFQENNLDLVRSHEVKDEGTRIEHDGKLITVFS 435

Qy 382 SITYATGGSNKGAYTILM-NOLMPHFTQY 408  
Db 436 PNVCDCOMGNGKGAIRFREAPDMKPNITP 463

RESULT 15  
US-10-424-599-191442  
; Sequence 191442, Application US/10424599  
; Publication No. US20031077A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J  
; APPLICANT: Kovacic, David K  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement

Query Match 16.2%; Score 557.5; DB 4; Length 533;  
Best Local Similarity 39.4%; Pred. No. 4.6e-37; Mismatches 109; Indels 31; Gaps 10;  
Matches 127; Conservative 55; MisMatches 109; Indels 31; Gaps 10;

Qy 94 AKIE-LPIRKNHIDLLDVERKGRGNHLPKVVALURBAAMSLKOPNITSPVSTAVSQ 152

FILE REFERENCE: 38-21(53223)B  
CURRENT APPLICATION NUMBER: US/10/424, 599  
CURRENT FILING DATE: 2003-04-28  
NUMBER OF SEQ ID NOS: 285664  
SEQ ID NO: 191442  
LENGTH: 533  
TYPE: PRT  
ORGANISM: Glycine max  
FEATURE:  
NAME/KEY: unsure  
LOCATION: (1)..482  
OTHER INFORMATION:  
OTHER INFORMATION: unsure at all Xaa locations  
OTHER INFORMATION: Clone ID: PAT\_MRT3847\_14893C.1.pep  
US-10-424-599-191442

Query Match 16.3%; Score 561; DB 4; Length 482;  
Best Local Similarity 38.4%; Pred. No. 2e-37; Mismatches 119; Indels 52; Gaps 10;  
Matches 135; Conservative 46; MisMatches 119; Indels 52; Gaps 10;

Qy 60 POAAGRNKNOYQGSANVSVLDDKDDLIVETEFGDIDVNAKELPIRKONHIDLLDVERKGRGNR 119  
Db 158 POGSGAR--IEGDA--VTLDFVKQKIEDF-----KNOKF 187

Qy 120 LHPKTVALLREAAKSLKQPNISPVSTAVSQVTVCGDJKHGKUDLILVHKGNSLSS 179  
Db 188 LHKRYAFPOVIMOTKRAQJLQPSLUDHVPDKHFTVCPDVQYQYDILNIFELNGLSPSEB 247

Qy 180 NPVYFNGDFVDRGKRGKGLBLVLLSLYLASPAVNFVLRGHRBDSYNARYGPIREYESKY 239  
Db 248 NPVYFNGDFVDRGSSLLEVLTIPAKCMPSATVLAGNHSKSKMKVGFGRGSRSKL 307

Qy 240 PRNKKRILAFIDBEVYVWPLGSVLMRVLVHGG-FS-DETSLDLIIKSIDRQKVSILRP 297  
Db 308 NETFVFLPA--EVPCCLPLAHVINBKVFVHGGLESVDGVKVSDIRSINKR----FCBEP 359

Qy 298 PLTDGSPFLDKTWRQOQIDMWSDPQMTGCPVNTIRAGCYWFGGPVDTNFTLQRHRLSYV 357  
Db 360 P-----BEGIMCLLWSDPROLPGRGSPSK-RGVLSFGADVTTRFLQENNLDLV 408

Qy 358 RSHECKENGHRFMHDNKKITTSASYATGGSNKGAYTILM-NOLMPHFTQY 408  
Db 409 RSHEVKDEGVEIHLHDKLITVFSAPNYCQDCOMGNGKAFFIRFREAPDLPKPNITP 460

RESULT 16  
US-10-298-638-28  
; Sequence 28, Application US/10298638  
; Publication No. US20030177527A1  
; GENERAL INFORMATION:  
; APPLICANT: HARING, Michel A. et al.  
; TITLE OF INVENTION: NUCLEOTIDE SEQUENCES CODING SIGNAL TRANSDUCTION  
; TITLE OF INVENTION: COMPONENTS IN DURABLE AND BROAD-RANGE RESISTANCE  
; TITLE OF INVENTION: STRATEGIES BASED ON PLANT DEFENCE  
; FILE REFERENCE: 2121-0167P  
; CURRENT APPLICATION NUMBER: US/10/298, 638  
; CURRENT FILING DATE: 2002-11-19  
; PRIORITY APPLICATION NUMBER: 10/298, 638  
; PRIORITY FILING DATE: 2002-11-19  
; NUMBER OF SEQ ID NOS: 28  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO: 28  
; LENGTH: 533  
; TYPE: PRT  
; ORGANISM: Arabidopsis thaliana  
; US-10-298-638-28

Db 214 ARFEGEVTIDPVKIMMEDPKNQK-TLHKRKYAQIVIQLTRQILALPSLVDSVPHKH 271  
 Qy 153 VTVCGDLHKGKLDLUVWVKNGLPSSSNPYVFGDFVDRGKRGLEVLILLSLYAPNA 212  
 Db 272 ITVCGDVHGQPYBLNTPFLINGLSPSEENPVLNGDFVDRGSFSVERILITLPAFKCMCPSS 331  
 Qy 213 VFLNRGNHEDSUMMARYGPRREVESKYPKRNRKILAFD--EVYRMPLGSVLSNSRLVLI 269  
 Db 332 IYLAQRGHESKSMSKTYGPRGRSKLSBK---FDLFAKVFCLPLAHVINGKPFV 385  
 Qy 270 VHGG-FS-DSTSILIKSIDRGKVASTLRPLPTDGEPLDKTENQOIDPIMSDPOTNGC 327  
 Db 386 VHGGLFSDVGKUSDIRADR-----FCEP-----BEGLMCRLLWSDPOPLGR 430  
 Qy 328 VPVTLRGAAGWFGDPVTNFQHLRSVRSRSHCKPGHFRPHDKNITISASNYAI 387  
 Db 431 GPRSK-RGVGLSGSGDVTKRLQDNNDLIVRSHEVKDGYEVHDGKLITVSAPNCQDQ 489  
 Qy 388 GSNKGAYVRLN-NOLMHPHQV 408  
 Db 490 MGKNGAFAFREAPDMKPNVTF 511

RESULT 17

US-10-369-493-22595  
 ; Sequence 22595, Application US/10369493  
 ; Publication No. US20030233675A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Cao, Yongwei  
 ; APPLICANT: Hinkle, Steven C.  
 ; APPLICANT: Slater, Steven C.  
 ; APPLICANT: Goldman, Barry S.  
 ; APPLICANT: Chen, Xianfeng  
 ; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF PLANTS WITH IMPROVED PROPERTIES  
 ; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF PLANTS WITH IMPROVED PROPERTIES  
 ; CURRENT FILING DATE: 2003-12-28  
 ; PRIORITY APPLICATION NUMBER: US/10/369, 493  
 ; PRIORITY FILING DATE: 2002-02-21  
 ; NUMBER OF SEQ ID NOS: 47374  
 ; SEQ ID NO: 4091  
 ; LENGTH: 598  
 ; TYPE: PRT  
 ; ORGANISM: Neurospora crassa  
 ; FEATURE:  
 ; NAME/KEY: unbure  
 ; LOCATION: (1)..(598)  
 ; OTHER INFORMATION: unsure at all Xaa locations  
 US-10-369-493-4091

Query Match 14.6%; Score 502.5; DB 4; Length 598;  
 Best Local Similarity 29.2%; Pred. No. 2e-32; Indels 141; Mismatches 55; Gaps 12;  
 Matches 141; Conservative 29.2%; Pred. No. 2e-32; Indels 139; Gaps 12;

Qy 12 IOKWYRHOARREMORRCRNQWIFQNLKASEDOAEYKPFENDLIKMPQAGRKNQIG 71  
 Db 158 VZAYYRBRATA-----YAATINPKRVRKDFKTCVKAP----- 189  
 Qy 72 SAHVSVLDDKD---DLV-----EFGDITVNAKTELPIRKHNHDLLID--- 110  
 Db 190 -----DNKDAKLUJVECKTVQRLAFFAAIEVGDELSAEGDYESMADSYQVRL 242  
 Qy 111 -----VFRKRGHRLPKVQVALTRRANKSLKQLPNITPPVSTAVSQVTVG 157  
 Db 243 EGENMTQEFIDMMIERFRKGKLUHKVYQVITAVRNIVNEMPTWVTDIPEDVQVLC 302

Query Match 15.1%; Score 519; DB 4; Length 473;  
 Best Local Similarity 35.2%; Pred. No. 5.9e-34;  
 Matches 119; Conservative 66; Mismatches 127; Indels 26; Gaps 8;

Qy 73 AHVSVLDDKDLYVEBPGDIVAKELPLTRKRNHDLLIDVFRKKGRNRLAHPKVALIREA 132  
 Db 139 ANNTED-MDPSDQCVI--LEROKITKEFEDMKEP--CQKGULPKRAYSILRD 191  
 Qy 133 AKSLKOLPNISPVSTAVSOQVWCGDLGKLDLUVWVKNGLPSSSNPYVFGDFVDRG 192  
 Db 192 KSELEKPTPSLIDIPVKGDETLVCGDTGQYFDLNLNPKLHGPPSPNPKYLNGDFVDRG 251  
 Qy 193 KQGLRVUILLSLYALPNAVFLNPGHEDSUMMARYGPRREVESKYPKRNRKILAFD 252  
 Db 252 SWSTEVAFTLAVKYLDPDAVPRNGHETDMDKVKYGPGRCSRKY---NERTENPSE 308  
 Qy 253 VWRMLPLASVLSVRLVHGG---DSTSILIKSIDRGKVASTLRPLPTDGEPLDKTEW 310  
 Db 309 TTSILPLASLISLVSYLVHGGFSDDNLTDLQRNIRPS---KQGPQSG----- 356  
 Qy 311 QQTFDIIMWSDPQTMGCVNPTLRAAGWFGDPVTNFQHLRSVRSRSHCKPNSHFBM 370  
 Db 357 -LMMEMLWTDQFAPGPGPSK-RGVGLQFGPDVSKRFBCBANGLKAVIRSHEVRDQGIEVE 414  
 Qy 371 HDKIKITIFSAASNYAAGSNKAYTRLNNOLMHPHQV 408

Db 415 HDGCVTIVFSAPNCDCSTGNLGAIVKVKEDMBLDFHQF 452  
 RESULT 18  
 US-10-369-493-4091  
 ; Sequence 4091, Application US/10369493  
 ; Publication No. US20030233675A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Cao, Yongwei  
 ; APPLICANT: Hinkle, Steven C.  
 ; APPLICANT: Slater, Steven C.  
 ; APPLICANT: Goldman, Barry S.  
 ; APPLICANT: Chen, Xianfeng  
 ; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF PLANTS WITH IMPROVED PROPERTIES  
 ; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF PLANTS WITH IMPROVED PROPERTIES  
 ; CURRENT FILING DATE: 2003-12-28  
 ; PRIORITY APPLICATION NUMBER: US/10/369, 493  
 ; PRIORITY FILING DATE: 2002-02-21  
 ; NUMBER OF SEQ ID NOS: 47374  
 ; SEQ ID NO: 4091  
 ; LENGTH: 598  
 ; TYPE: PRT  
 ; ORGANISM: Neurospora crassa  
 ; FEATURE:  
 ; NAME/KEY: unbure  
 ; LOCATION: (1)..(598)  
 ; OTHER INFORMATION: unsure at all Xaa locations  
 US-10-369-493-4091

Query Match 14.6%; Score 502.5; DB 4; Length 598;  
 Best Local Similarity 29.2%; Pred. No. 2e-32; Indels 141; Mismatches 55; Gaps 12;  
 Matches 141; Conservative 29.2%; Pred. No. 2e-32; Indels 139; Gaps 12;

Qy 12 IOKWYRHOARREMORRCRNQWIFQNLKASEDOAEYKPFENDLIKMPQAGRKNQIG 71  
 Db 158 VZAYYRBRATA-----YAATINPKRVRKDFKTCVKAP----- 189  
 Qy 72 SAHVSVLDDKD---DLV-----EFGDITVNAKTELPIRKHNHDLLID--- 110  
 Db 190 -----DNKDAKLUJVECKTVQRLAFFAAIEVGDELSAEGDYESMADSYQVRL 242  
 Qy 111 -----VFRKRGHRLPKVQVALTRRANKSLKQLPNITPPVSTAVSQVTVG 157  
 Db 243 EGENMTQEFIDMMIERFRKGKLUHKVYQVITAVRNIVNEMPTWVTDIPEDVQVLC 302

Query Match 15.1%; Score 519; DB 4; Length 473;  
 Best Local Similarity 35.2%; Pred. No. 5.9e-34;  
 Matches 119; Conservative 66; Mismatches 127; Indels 26; Gaps 8;

Qy 73 AHVSVLDDKDLYVEBPGDIVAKELPLTRKRNHDLLIDVFRKKGRNRLAHPKVALIREA 132  
 Db 139 ANNTED-MDPSDQCVI--LEROKITKEFEDMKEP--CQKGULPKRAYSILRD 191  
 Qy 133 AKSLKOLPNISPVSTAVSOQVWCGDLGKLDLUVWVKNGLPSSSNPYVFGDFVDRG 192  
 Db 192 KSELEKPTPSLIDIPVKGDETLVCGDTGQYFDLNLNPKLHGPPSPNPKYLNGDFVDRG 251  
 Qy 193 KQGLRVUILLSLYALPNAVFLNPGHEDSUMMARYGPRREVESKYPKRNRKILAFD 252  
 Db 252 SWSTEVAFTLAVKYLDPDAVPRNGHETDMDKVKYGPGRCSRKY---NERTENPSE 308  
 Qy 253 VWRMLPLASVLSVRLVHGG---DSTSILIKSIDRGKVASTLRPLPTDGEPLDKTEW 310  
 Db 309 TTSILPLASLISLVSYLVHGGFSDDNLTDLQRNIRPS---KQGPQSG----- 356  
 Qy 311 QQTFDIIMWSDPQTMGCVNPTLRAAGWFGDPVTNFQHLRSVRSRSHCKPNSHFBM 370  
 Db 357 -LMMEMLWTDQFAPGPGPSK-RGVGLQFGPDVSKRFBCBANGLKAVIRSHEVRDQGIEVE 414  
 Qy 371 HDKIKITIFSAASNYAAGSNKAYTRLNNOLMHPHQV 408

Db 585 SQF 587

RESULT 19

US-10-450-763-40347  
; Sequence 40347, Application US/10450763  
; Publication No. US2005019675A1  
; GENERAL INFORMATION:  
; APPLICANT: Hyseq, Inc.  
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES  
; FILE REFERENCE: 790CIP/US  
; CURRENT APPLICATION NUMBER: US/10/450-763  
; CURRENT FILING DATE: 2003-06-11  
; PRIORITY APPLICATION NUMBER: PCT/US01/08631  
; PRIORITY FILING DATE: 2001-03-30  
; PRIORITY APPLICATION NUMBER: 09/540,217  
; PRIORITY FILING DATE: 2000-03-31  
; PRIORITY APPLICATION NUMBER: 07/649,167  
; PRIORITY FILING DATE: 2000-08-23  
; NUMBER OF SEQ ID NOS: 60736  
; SOFTWARE: Custom  
; SEQ ID NO: 40347  
; LENGTH: 714  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURES:  
; NAME/KEY: DOMAIN  
; LOCATION: (384)..(3856)  
; OTHER INFORMATION: SERINE/THREONINE PHOSPHATASE FAMILY SIGNATURE domain  
; OTHER INFORMATION: identified by eMATRIX, accession number PR00114A, p-value=4.600e-  
; FEATURE:  
; NAME/KEY: DOMAIN  
; LOCATION: (321)..(696)  
; OTHER INFORMATION: Ser/Thr protein phosphatase domain identified by Pfam,  
; OTHER INFORMATION: accession name Sphosphatase, E-value=4.6e-92, Ppm Score of 319.  
; US-10-450-763-40347

Query Match 14.4%; Score 496.5; DB 5; Length 714;  
Best Local Similarity 29.2%; Pred. No. 8e-32; Mismatches 147; Indels 131; Gaps 15;  
Matches 140; Conservative 62; Mismatches 147; Indels 131; Gaps 15;

Qy 41 SRQDQABEELYKFFNDLKHMPQAAGRKNQYO-----GSAH----SVLDDKD- 82  
Db 231 SEBHAATIYL--VVVKPKHDKAKKYKBCNKKIVQKAKBRAIADEDEHRSVVDSDLI 287

Qy 83 --DLVERPG--DIVNAKIELPIRKONIDLIDVFRERGRGRHLPKVVALTRRAKSIK 137  
Db 288 EMTIEDBYSGPKPLEDKGKVTSFMK-----JMWQYDOK-KLHRKCAYQILVQKEVL 341

Qy 138 QPNISPVSTAVSQQTIVCGDQGLKGKDDLLVVLVHKNGLPSSNPV----- 183  
Db 342 KUSTLVEETKTEKITEKIVTGCQDTHQFVDTLNIFELNGLPSITNPVFSAPPSPHLHQ 401

Qy 184 -----FNQDFVDRGKRGKLEVLVLLLSLYTAF 209  
Db 402 PQQVQCTEWBRPLPGKOTVLERGLAMLGHQOQIFNGDFVQGSFSEVILTLFGKULY 461

Qy 210 PNAVFLPLRGNHEDSVMARYGTREVSKYPRKRKLAFTDDEVYKWLPLQASVLSVRLY 269  
Db 462 PDPFHLLRGNHEDTNMOMIQYGFGEVAKYT---ADMELPSERVESWLPQACINGKVL 518

Qy 270 VIGGP--SDSTSIDLINSIDRGKTVSLRPPITGSPGLDKEWQDFPIMMSDPO---- 322  
Db 519 MHGGLFSDBGVTLDDIRKTRR----OPP--DSALPRRMGCFVPLARNNSSEPPQPHLR 571

Qy 323 -----ATMGCCVPNTL-----RGACWFWGPDVNTNFQ 349  
Db 572 LQPVHLLGAVMASTVUTMVRGEGTVTAEHTCPYLPFGRSISKRGVSCQFDPDVTKAFLE 631

Qy 350 RHLSYVYVIRSHCKPQNGHEFMNDKLTITFSASNYAAGSNGAYRL-NQHLMPEFVQY 408  
Db 632 ENNDYIIRSHEVRAEGIEVAHGGRCVTVFAPNCYDOMGMKASYIHLQGDRLPQPHQF 691

---

RESULT 20  
; US-10-369-493-21980

Query Match 13.9%; Score 479.5; DB 4; Length 513;  
Best Local Similarity 32.3%; Pred. No. 1.2e-30; M1-matches 154; Indels 51; Gaps 13;  
Matches 130; Conservative 32; Mismatches 154; Indels 51; Gaps 13;

Qy 2 DENATAIAFIQKWTTRHARREM--QRRCNWQIIFONLEYASBODQAELEYKFENDLIK 58  
Db 113 DPAATKALLTCDRFRERKRAIGGAENEAKISLCQTLNLSFDANADANY----- 165

Qy 59 MPQAAGRKNQYQGSAHVSVDKODLVEERGDIWAKLRPIRKOHIDLIDVPRKGRN 118  
Db 166 ----BQPKLFE----OLYTDKN-----AFKAKKNSOFTSKVMDNLK--GK 207

Qy 119 RLHPKVKVALTIREAKSLKQLPNISPV--STAVSQQVTWGQDHLGKDDLLVVLKNGLP 176  
Db 208 YLPKKVVAITSHADTFLRQPSPMTELENSTPDYKISVGQDTHQFQFVDTNLPLPKGV 267

Qy 177 SSNNPVTYFNGDPYDVRGKRGAVLULLLSSLYTAFPNVAVLNGRNHEDSVMARYGTRIVE 236  
Db 328 YTKY--SORIENMFQAFSPESFESPLATINNDYLVMHGGPLPSATSLSDFKNIDR----F 379

Db 268 GPKHTYLFNGDPVDRGWSCEVALLFYCLKLHPPNNFLRGNHEDSNMNKLYGPFEDCK 327

Qy 237 SKYPRHKRKLAFIPLVYRMLPLGSVLSVNLVIGGG-SD-SISLDLKSIDRKVSI 294  
Db 328 YTKY--SORIENMFQAFSPESFESPLATINNDYLVMHGGPLPSATSLSDFKNIDR----F 379

Qy 295 LRPPUDTGSPDKEWMQOIDMWSDPQATMGCVNTLRLGAGVWIGPDPVTDNFLORHLS 354  
Db 380 AQPP-RDG-----AFHLLWADPQRANGMGQS--RGIHGATGDPDITDFRANKLRL 428

Qy 355 YVTRSHBCKPQNGHEFMNDKLTITFSASNYAAGSNGAYRL 397  
Db 429 KIFRSHLRCMGVQFQKGKLMIVFSAPNCYCDSONGLGGVHV 471

---

RESULT 21  
; US-11-097-143-5988

Sequence 5988, Application US/11097143  
; Publication No. US20050208558A1  
; GENERAL INFORMATION:  
; APPLICANT: Venter, J. Craig  
; TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID  
; TITLE OF INVENTION: ARRAYS FOR DETECTING EXPRESSION OF 10,000 OR MORE  
; TITLE OF INVENTION: DROSOPHILA GENES.  
; FILE REFERENCE: C1000728

CURRENT APPLICATION NUMBER: US/11/097,143  
; CURRENT FILING DATE: 2005-04-04  
; PRIORITY APPLICATION NUMBER: 60/157,832  
; PRIORITY FILING DATE: 1999-10-05  
; PRIORITY APPLICATION NUMBER: 60/160,191  
; PRIORITY FILING DATE: 1999-10-19  
; PRIORITY APPLICATION NUMBER: 60/161,932

RESULT 20  
; US-10-369-493-21980

PRIOR FILING DATE: 1999-10-28  
; PRIOR APPLICATION NUMBER: 60/164,769  
; PRIOR FILING DATE: 1999-11-12  
; PRIOR APPLICATION NUMBER: 60/173,383  
; PRIOR FILING DATE: 1999-12-28  
; PRIOR APPLICATION NUMBER: 60/175,693  
; PRIOR FILING DATE: 2000-01-12  
; PRIOR APPLICATION NUMBER: 60/184,831  
; PRIOR FILING DATE: 2000-02-24  
; PRIOR APPLICATION NUMBER: 60/191,637  
; PRIOR FILING DATE: 2000-03-23  
; NUMBER OF SEQ ID NOS: 43008  
; SOFTWARE: FASTSEQ for Windows Version 4.0  
; SEQ ID NO: 5988  
; LENGTH: 346  
; TYPE: PRT  
; ORGANISM: DROSOPHILA  
US-11-097-143-5988

Query Match 13.4%; Score 460.5; DB 6; Length 346;  
Best Local Similarity 33.8%; Pred. No. 2.6e-29;  
Matches 104; Conservative 63; Mismatches 120; Indels 21; Gaps 6;

QY 101 RKKHHIDLLIDVFRKKR-GNRLHPKVYKVALITREAAKSLLKOLPNTPSPVSTAVSQQTVWCGNL 159  
Db 21 RMSQQLDVIGVQKTMAGVNRRAGNISAEATITYCQASRFLSQQMLBLASPVRICGDL 80  
Qy 160 HAKLDDILIVLHLNGLPPSSSNPVPNGFVDPDKRGKGLLEVLLSLYLAPPNAVLNRGN 219  
Db 81 HSGPKDILRIFFOCGVGPPLSN-TLFGLGVDRGHSISRTSLUTYKURPYEPFLLAGN 139  
Qy 220 HEHSVMMARYGFTREVSKYPRHRKRLAFIDBVRWLPGLSVLNRSVLTVGFFS-DST 278  
Db 140 HESADALNRYVGFDECKRY-SIKLWRSFD-CYDCMVAIAIDRIFCVHGGSLSPDLN 196  
Qy 279 SLDLKSIDRKGKVTSILRPLTDGEPBLKTENQDFDINWSDFQATMGCVPTNTGAGGW 338  
Db 197 NLDDIRRLNR-----PTDVSQDGCLLCDLWSDPDTETGTWASNDRGVSPT 241  
Qy 339 RGDVDTONPLQRHLSYVIRSHECKPNHGFMDMKITIPSAASNYAIGSNKGAYIRLN 398  
Db 242 FGANTIVSGFLMOKFNLTIVRAHOWEDGEYEFRADQLVTFSPNCFDONGAVAWD 301  
Qy 399 NQLMPHFV 406  
Db 302 AKLVCHFV 309

RESULT 23  
US-10-425-115-225312  
; Sequence 25312; Application US/10425115  
; Publication No. US20040214272A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants  
; FILE REFERENCE: 38-21(5322)B  
; CURRENT APPLICATION NUMBER: US/10/425,115  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 369326  
; SEQ ID NO: 225312  
; LENGTH: 443  
; TYPE: PRT  
; ORGANISM: Zea mays  
; FEATURE:  
; OTHER INFORMATION: Clone ID: MRT4577\_137074C.1.pep  
US-10-425-115-225312

Query Match 13.0%; Score 447.5; DB 4; Length 443;  
Best Local Similarity 33.5%; Pred. No. 4.5e-28;  
Matches 113; Conservative 45; Mismatches 104; Indels 75; Gaps 8;

Qy 120 LHPKVYKVALITREAAKSLLKOLPNTPSPVSTAVSQQTVWCGNLVKGKDDLLVWAKHNGLPPSS 179  
Db 74 LPPALVQRIUAAAMAIHLRPPFLVWDPQCGAVVNGDHWQQLHNVPIRDAAGPSED 133  
Qy 180 NPYVENGDFVDRGKGLLEVLLSLYLAPPNAVLNRGNEDSYNARYGFTREVSKY 239  
Db 134 RIFVFNGDVVRGAWGLETFLLLAWKVLNCVFLURGNHESKCVTSVVGPEKVMY 193  
Qy 240 ---PRMHETLARDBVTVWLPGLSVLNRSVLTVGFFEDTSIDLISDRKG---- 290  
Db 194 KDGQPIQYKRFKLFRCED---LPLATIAGCYVTAHGGIFTGFTIVUPSKSKKKGKKGKY 249  
Qy 291 ---YVSLRPP---LTDGBPLDKTEWQOIFDIMSDPOA 323  
Db 250 KANFTWDJISMKIGSLDELJMKARRTWDPMWEGSNLPG----DVLNSDPL 297  
Qy 324 TMGCWPTNTLARGAGWFGPDVTDFNORHLSYVIRSHECKPNHGFMD---- 372  
Db 298 QMGLSPNPKERGIGLWGPDTIQFLRNHLKLIRSHE-GFDARDKRRHOLGMKDGYTID 356  
Qv 373 ---NKITITPSASNYAIG----SNKGAYIRLN 398

Db 357 HEVECGKULITFSAPDYPQFQASEERYNNGAYIVLS 393 ; NUMBER OF SEQ ID NOS: 73128  
; SEQ ID NO: 50385  
; LENGTH: 339  
; TYPE: PRT  
; ORGANISM: zea maya  
; FEATURES:  
; OTHER INFORMATION: Clone ID: LIB3062-002-C7\_FLT.pep  
; US-10-425-114-50385  
Query Match 12.6%; Score 432.5; DB 4; Length 326;  
Best Local Similarity 30.3%; Pred. No. 4.9e-21;  
Matches 105; Conservative 69; Mismatches 125; Indels 47; Gaps 9;  
Db 1 LRHTPRSPGRK--EGGKOMMR----APMGMEGAADMTR----LVEGGRG 46  
Qy 76 SVADDKDDOLVERSGDIVNAKIELPIRKRNHIDLLIDVFRKKRGNRLHRYVALLRBEAKS 135  
Db 4 SVLD-----DINRLLSVTRPQ-----KVQLSBSERIROLCVSREI--- 41  
Qy 136 LKQLPNTSPVSTAVSQYTCVDLHGGLDLWHLHNGLOSSSSNPYVNGDPYDVKRG 195  
Db 42 FLQOPNLRLRAPK---ICGDVHGOYSDLRLPBYGGLPPRAN-YPLGlyvdRQGS 96  
; PRIORITY: 38-21(5323)B  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: CURRENT APPLICATION NUMBER: US/10/425,114  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO: 254277  
; LENGTH: 326  
; TYPE: PRT  
; ORGANISM: Glycine max  
; FEATURES:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_71636C.1.pep  
; US-10-424-599-254277  
Query Match 12.7%; Score 435.5; DB 4; Length 323;  
Best Local Similarity 31.2%; Pred. No. 2.7e-27;  
Matches 104; Conservative 66; Mismatches 124; Indels 39; Gaps 9;  
Db 8 DINRLLEVTPRAKQVOLSSHRO-----LCVSRBI--FLQQPMLLEAP 54  
Qy 149 VSGQVTVGCDLHGKDDLVVHLHNGLOSSSSNPYVNGDPYDVKRGKGLLVLVLLSYLA 208  
Db 55 IK---ICGDVHGOYSDLRLPBYGGLPPRAN-YPLGlyvdRQGS 109  
Qy 209 FPNAVPLRGNHEDSVMMARPGTIREVKPSVSKYPRNHKRILAFIDEVYRMLPLGSVLNSRVL 268  
Db 110 YPENPFPLRGNHHECASINRIGYFDECKERF--NVRUWKTPDFCNCLPVALVDEKL 166  
Qy 269 IVHCGFS-DSTSIDLKISRDKGVYSLIRPLPTDGERPLDKTSMQFIDIMWSDPOATMGC 327  
Db 167 CMIGGLSPDNNLQDQIRNQRL-----PAVPDTGCLLMSDPSPKDVOG 211  
Qy 328 VPNTLTRGLAGWWSGPDVDTNFLORRHLSTVTSRSHCKPQHBFEMHDNKLTISASNTYA 387  
Db 212 WGMDNRGVSYTFGADKVQSQQFLQRDLIVCARHOVEDGEYFPANCOLVTIFSAPNYCGE 271  
Qy 388 GSWKGAYIRLNQMLMPHYQYQISASQTKRLSP 420  
Db 272 FDNAGAMMSVDETMCSF-QILKPADKKAKLNF 303  
; RESULT 25  
US-10-425-114-50385 ; Sequence 254277, Application US/10424599  
; Publication No. US20040034888A1 ; General Information:  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J ; APPLICANT: Kovacic David K  
; APPLICANT: Zhou Yihua ; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; TITER OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(5323)B ; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO: 254277  
; LENGTH: 326  
; TYPE: PRT  
; ORGANISM: Glycine max  
; FEATURES:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_71636C.1.pep  
; US-10-424-599-254277  
Query Match 12.6%; Score 432.5; DB 4; Length 326;  
Best Local Similarity 30.3%; Pred. No. 4.9e-21;  
Matches 105; Conservative 69; Mismatches 125; Indels 47; Gaps 9;  
Db 1 LRHTPRSPGRK--EGGKOMMR----APMGMEGAADMTR----LVEGGRG 46  
Qy 76 SVADDKDDOLVERSGDIVNAKIELPIRKRNHIDLLIDVFRKKRGNRLHRYVALLRBEAKS 135  
Db 4 SVLD-----DINRLLSVTRPQ-----KVQLSBSERIROLCVSREI--- 41  
Qy 136 LKQLPNTSPVSTAVSQYTCVDLHGGLDLWHLHNGLOSSSSNPYVNGDPYDVKRG 195  
Db 42 FLQOPNLRLRAPK---ICGDVHGOYSDLRLPBYGGLPPRAN-YPLGlyvdRQGS 96

QY 196 IELVILLISLYLAEPNAVINLARGNHEDSYMMARYGPIREVSKYPRNKRILAPIDEVYR 255 ; Sequence 9917, Application US/10739930  
Db 97 LETCILULLAYKIKYPENPLTFLRGHECASINRYGPIVNECKRPF--NVRWKTFDGFN 153 ; Publication No. US20040216190A1  
; GENERAL INFORMATION:  
; APPLICANT: Kovacic, David K.  
; TITLE OF INVENTION: NUCLEAR ACID MOLECULES AND OTHER MOLECULES ASSOCIATED WITH FILE REFERENCE: 38-21(53377)B  
; TITLE OF INVENTION: PLANTS AND USES THEREOF FOR PLANT IMPROVEMENT CURRENT APPLICATION NUMBER: US/10/739,930  
; CURRENT FILING DATE: 2003-12-18  
; NUMBER OF SEQ ID NOS: 11088  
; SEQ ID NO 9917  
; LENGTH: 325  
; TYPE: PRT  
; ORGANISM: Triticum aestivum  
; FEATURES:  
; OTHER INFORMATION: Clone ID: TRIAR-23APR03-C212\_73.p

## RESULT 27

US-10-425-114-65099  
; Sequence 65099, Application US/10425114

; Publication No. US20040034888A1

; GENERAL INFORMATION:

; APPLICANT: Liu, Jingdong  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Kovacic, David K.  
; APPLICANT: Screen, Steven E  
; APPLICANT: Tabaska, Jack E  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With FILE REFERENCE: 38-21(53313)B  
; CURRENT APPLICATION NUMBER: US/10/425,114  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 73128  
; SEQ ID NO 65099  
; LENGTH: 380  
; TYPE: PRT  
; ORGANISM: Glycine max

; OTHER INFORMATION: Clone ID: 700833616\_FLI.pep

; OTHER INFORMATION: Clone ID: US-10-425-114-65099

Query Match 12.5%; Score 430.5; DB 5; length 325;  
Best Local Similarity 31.1%; Pred. No. 7.2e-27; Matches 99; Conservative 68; Mismatches 122; Indels 29; Gaps 6;  
Matches 99; Conservative 68; Mismatches 122; Indels 29; Gaps 6;

QY 89 GDTWNAKELPLPIRKHNDLIDVPLRKKGRRNLRLPKVAVILRRAKSLSKLQQLPNISPVSTA 148 ; Sequence 9917, Application US/10739930  
Db 9 GAMEGAALDEVVRR--LVEGRGRGQVQSAEIRQLCVDAAKRVPSOPNL--LR 59 ; Publication No. US20040216190A1  
; GENERAL INFORMATION:  
; APPLICANT: Liu, Jingdong  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Kovacic, David K.  
; APPLICANT: Screen, Steven E  
; APPLICANT: Tabaska, Jack E  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With FILE REFERENCE: 38-21(53377)B  
; CURRENT APPLICATION NUMBER: US/10/739,930  
; CURRENT FILING DATE: 2003-12-18  
; NUMBER OF SEQ ID NOS: 11088  
; SEQ ID NO 9917  
; LENGTH: 325  
; TYPE: PRT  
; ORGANISM: Triticum aestivum  
; FEATURES:  
; OTHER INFORMATION: Clone ID: TRIAR-23APR03-C212\_73.p

## US-10-739-930-9917

Query Match 12.5%; Score 430.5; DB 5; length 325;  
Best Local Similarity 31.1%; Pred. No. 7.2e-27; Matches 99; Conservative 68; Mismatches 122; Indels 29; Gaps 6;

; Sequence 9917, Application US/10739930

; Publication No. US20040216190A1

; GENERAL INFORMATION:

; APPLICANT: Liu, Jingdong  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Kovacic, David K.  
; APPLICANT: Screen, Steven E  
; APPLICANT: Tabaska, Jack E  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With FILE REFERENCE: 38-21(53313)B  
; CURRENT APPLICATION NUMBER: US/10/425,114  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 73128  
; SEQ ID NO 64475  
; LENGTH: 334  
; TYPE: PRT  
; ORGANISM: Zea maya  
; FEATURES:  
; OTHER INFORMATION: Clone ID: LIB3969-010-E2\_FLI.pep

## US-10-425-114-64475

Query Match 12.5%; Score 430.5; DB 4; length 334;

; Sequence 64475, Application US/10425114

; Publication No. US20040034888A1

; GENERAL INFORMATION:

; APPLICANT: Liu, Jingdong  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Kovacic, David K.  
; APPLICANT: Screen, Steven E  
; APPLICANT: Tabaska, Jack E  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With FILE REFERENCE: 38-21(53313)B  
; CURRENT APPLICATION NUMBER: US/10/425,114  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 73128  
; SEQ ID NO 64475  
; LENGTH: 334  
; TYPE: PRT  
; ORGANISM: Zea maya  
; FEATURES:  
; OTHER INFORMATION: Clone ID: LIB3969-010-E2\_FLI.pep

## RESULT 28

US-10-739-930-9917  
; Sequence 9917, Application US/10739930

; Publication No. US20040216190A1

; GENERAL INFORMATION:

; APPLICANT: Liu, Jingdong  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Kovacic, David K.  
; APPLICANT: Screen, Steven E  
; APPLICANT: Tabaska, Jack E  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With FILE REFERENCE: 38-21(53313)B  
; CURRENT APPLICATION NUMBER: US/10/425,114  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 73128  
; SEQ ID NO 64475  
; LENGTH: 334  
; TYPE: PRT  
; ORGANISM: Zea maya  
; FEATURES:  
; OTHER INFORMATION: Clone ID: LIB3969-010-E2\_FLI.pep

## RESULT 29

US-10-425-114-64475  
; Sequence 64475, Application US/10425114

; Publication No. US20040034888A1

; GENERAL INFORMATION:

; APPLICANT: Liu, Jingdong  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Kovacic, David K.  
; APPLICANT: Screen, Steven E  
; APPLICANT: Tabaska, Jack E  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With FILE REFERENCE: 38-21(53313)B  
; CURRENT APPLICATION NUMBER: US/10/425,114  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 73128  
; SEQ ID NO 64475  
; LENGTH: 334  
; TYPE: PRT  
; ORGANISM: Zea maya  
; FEATURES:  
; OTHER INFORMATION: Clone ID: LIB3969-010-E2\_FLI.pep

Best Local Similarity 30.6%; Pred. No. 7.5e-27;  
 Matches 113; Conservative 67; Mismatches 134; Indels 55; Gaps 13;

Qy 60 POAAGRKNQYQGSARVSVLDDKDLVERFGDIVAK--IELPKRNHDIDLIVFRKRG 117  
 Db 16 PAAGG--OGGGIDAALLI--DDIRULEVRAPK-KQVQSESEFRCQTVSR--- 65

Qy 118 NRLHPKVYALLRREAKSILQLPNISPVSTAVSQVTWCGDAGKLDLVLVLAHNGLPS 177  
 Db 66 -----AIFLSQ-----PNLEAPK---ICGDINGQYSQDILRLFEYGGPP 105

Qy 178 SSNPYVENGDFYDRGKRGLEVLILLSLYTAFFNAFLNRGHEDSYNARIGFIREVES 237  
 Db 106 EAN-YLFGLGYDVRKGQSITCILLLAVKXKYPENFFLRGKVNRCASINRIGFYDECKR 164

Qy 238 KYNRHKRILAFIDEVRLPLGSVLSNRVLVHGGS-DSTSIDLIKSIDRKVYSLR 296  
 Db 165 RF---NVLWKVFTEGENTLPVAALIDKILCMHGSGSPDHLDSIKTOR----- 213

Qy 297 PPLTDGERLDKTMQQTIDIMSDP-QATMGCVPNTLRGAGWWFGPDVTDNLQRHLSY 355  
 Db 214 -----PTDVGQGLCDLMSDPGKDVOGMGND-RGVSTYFGADKVSBLFLORHIDL 265

Qy 356 VIRSHCKPKNSHFRMHDNKITITFSASHNYAIGSNKGAYTRNLNOLMPHFTYISASOT 415  
 Db 266 ICRAHOVVEDGYEFFADQLQVTFSAPNYCFCFDNAGAMMSVDETLMCSF-OILKPAERK 324

Qy 416 KRISPKORM 424  
 Db 325 QIYGAKQNV 333

RESULT 30

US-10-425-114-7017  
 ; Sequence 70117, Application US/10425114  
 ; Publication No. US20040034888A1

; GENERAL INFORMATION:

; APPLICANT: Liu, Jingdong  
 ; APPLICANT: Zhou, Yihua  
 ; APPLICANT: Kovacic, David K.  
 ; APPLICANT: Screen, Steven E  
 ; APPLICANT: Tabaska, Jack E  
 ; APPLICANT: Cao, Yongwei  
 ; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With Title of Invention: Plants and Uses Thereof for Plant Improvement  
 ; FILE REFERENCE: 38-21(53313)B  
 ; CURRENT APPLICATION NUMBER: US/10/425,114  
 ; CURRENT FILING DATE: 2003-12-18  
 ; NUMBER OF SEQ ID NOS: 11088  
 ; SEQ ID NO: 7415  
 ; LENGTH: 325  
 ; ORGANISM: Zea maya  
 ; TYPE: PRT  
 ; FEATURE:  
 ; OTHER INFORMATION: Clone ID: ZEAMA-23APRO3-C14584\_1.p  
 ; US-10-739-330-715

Query Match 12.5%; Score 429.5; DB 5; Length 325;  
 Best Local Similarity 31.6%; Pred. No. 8.7e-27;  
 Matches 99; Conservative 67; Mismatches 118; Indels 29; Gaps 6;

Qy 94 AKIHLPRKRNHDDILIVFRKRGNRHLPKVYLRLRREAKSILQLPNISPVSTAVSQV 153  
 Db 14 AADEVVR----LVGGGRQRQVQNSABIRQLCVAKOVLQSNL---LRIFAV 64

Qy 154 TWCQDPLAKDLDLVLVHNGLPSSSNPyENGDFYDRGKRGLEVLILLSLYTAFFNA 213  
 Db 65 KICGDINGQFVNLRLDQGKPTPS-YIYFGLGYDVRKGQSITCILLYAKVYDNT 123

Qy 214 FLNRGHEDSVNARIGFIREVESKPRNHRILAFIDEVRLPLGSVLSNRVLVHG 273  
 Db 124 YLRRGNHEDAKINRVYGFYDECKR---NVLWKVFTEGENTLPVAALIDKILCMHG 180

Qy 274 FS-DSTSIDLIKSIDRKVYVSLRPPTDGERLDKTYERQQLFDIMSDPQATMGCVPNTL 332  
 Db 181 LSPELTSIDLQKIER-----PTEDIPYGLCDLMSDPHSDTEGMBSD 225

Qy 333 RAGQWGPDVTDNLQRHLSYVIRSHCKPKNSHFRMHDNKITITFSASHNYAIGSNKG 392  
 Db 226 RGVSTYFGADKVSBLFLORHIDL 265

Qy 393 AVTRLNOLMPHFTYISASOT 405  
 Db 286 ALISIDSILMCSF 298

RESULT 32

US-10-739-930-9330  
 ; Sequence 9330, Application US/10739930  
 ; Publication No. US2004016190A1

; GENERAL INFORMATION:  
 ; APPLICANT: Kovacic, David K.  
 ; TITLE OF INVENTION: NUCLEIC ACID MOLECULES AND OTHER MOLECULES ASSOCIATED WITH



RESULT 35  
US-10-437-963-104697  
; Sequence 104697, Application US/10437963  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovacic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Wu, Wei  
; APPLICANT: Boukharov, Andrey A.  
; APPLICANT: Li, Ping  
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With FILE REFERENCE: 38-21(5322)B  
; CURRENT APPLICATION NUMBER: US/10/437-963  
; CURRENT FILING DATE: 2003-02-28  
; PRIORITY APPLICATION NUMBER: US 60/3160, 039  
; PRIORITY FILING DATE: 2002-02-21  
; NUMBER OF SEQ ID NOS: 47374  
; SEQ ID NO 2574  
; LENGTH: 316  
; TYPE: PRT  
; ORGANISM: Schizosaccharomyces pombe  
; FEATURE: NAME/KEY: unsure  
; LOCATION: (1)..(316)  
; OTHER INFORMATION: unsure at all Xaa locations

US-10-437-963-104697  
; Query Match 12.4%; Score 428.5; DB 4; Length 428;  
; Best Local Similarity 33.2%; Pred. No. 1. 6e-26; Matches 120; Conservative 48; Mismatches 134; Indels 59; Gaps 9;  
; OTHER INFORMATION: Clone ID: PAT\_MRT4530\_102005C.1.pep

QY 84 LYEEFGDVINAKELPIRKRNHIDLLVYFRKURGNRHNPKVALINREAAKSILKOLPNIS 143  
Db 32 LREWVAGLTTADWCWSRLSPADLPV-----LPAALVORLYAASVTIHLRSPNLV 83  
Qy 144 PYSTAVSQVTGCGDLHQKLDDLVVUHKGIPSSSNPVSFQDFDRKGKQGLEVLILL 203  
Db 84 RVDPRPDAVVVVGWQHQLHVITDAGEPSDDRVEFVGVDYDVRGANGLETTILL 143  
Qy 204 SIYLAFPVNAVFLNRGRNEDSVKARVGPIRETSKY---PRNHKRILAPIDEVYRNLPL 259  
Db 144 AWKVLPLNSVFLLRGNGHESKCYTIVSGFQEBOEVNFVKYKGQGPOVYRKLRCED---LPL 199  
Qy 260 GSTVNLNSVLVLLVIGGGFSSTSOLIKSDRG----KVSLRPPJGEPLDK---- 307  
Db 200 ANTIAGCIVTAHGGIFRGTIVMPSKRTRKGKKGKNGKQANSSDPTLMLGSLDELLKARR 259  
Qy 308 ---TEWQOI---FDIMSDPQATMGCVPUVNLGRAGWFMFGDVNTFLORHRLSIVIRS 359  
Db 260 TWLDPPMFGSNLIPGVLWSDLSLKGSLSSNKERGIGLUNGFDITQFLYTNLKAIRS 319  
Qy 360 HICKPNCHBFMED-----NKITITRSASNAYAIG----SNKGAYIRL 397  
Db 320 HB-GPDARDKRDLLGMNDGYTIDHHVCEGKLUTLISAPDYPQFOASEDRVNCGAYIVL 378  
Qy 398 N 398  
Db 379 N 379

RESULT 36  
US-10-369-493-2574  
; Sequence 2574, Application US/10369493  
; GENERAL INFORMATION:  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Hinkle, Gregory J.  
; APPLICANT: Slatter, Steven C.  
; APPLICANT: Goldin, Barry S.

RESULT 37  
US-10-425-115-233335  
; Sequence 233335, Application US/10425115  
; Publication No. US20040214272A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovacic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With FILE REFERENCE: 38-21(5322)B  
; CURRENT APPLICATION NUMBER: US/10/425,115  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 369326  
; SEQ ID NO 293335  
; LENGTH: 322  
; TYPE: PRT  
; ORGANISM: Zea mays  
; FEATURE: NAME/KEY: unsure  
; LOCATION: (1)..(322)  
; OTHER INFORMATION: unsure at all Xaa locations  
; FEATURE: OTHER INFORMATION: Clone ID: MRT4577\_30602C.1.pep

US-10-425-115-233335  
; Query Match 12.4%; Score 427.5; DB 4; Length 322;  
; Best Local Similarity 30.4%; Pred. No. 1. 3e-26;

Matches 112; Conservative 67; Mismatches 135; Indels 55; Gaps 13;

Qy 60 PQAAGKRNQYQGSARHVSVDKODLVBERGDIYNK--IELPKRNKHDLIDVPRKGK 117  
Db 5 PAAGC--QGGGIDAAU---DPIRRLKEVRTPRGKRVOLSEBIRQLCTVSR--- 54

Qy 118 NRLHPKVALLRRAAKSLKQLENISPVSTAVSQVTQCGDLHGKKDDLLVLVHNGLPS 177  
Db 55 -----AFLSQ-----PMLERAPIK---I-ODGIRDQYSQSLRLPEYGFPP 94

Qy 178 SSNPYVFNQDFVDRGKRGLEVLILLISLYLAPPNAFLNRGNHEDSVNARYGPIRES 237  
Db 95 EAN-YLFGLDYVDRGKQSLETICLILAYKIKYPENFLLRGHECASINRYGF 153

Qy 238 KPYRNHKILAF1DEYVWPLIGSVLNRLVINGGFS-DTSLDLKSIDRQKVSLR 296  
Db 154 RP---NVRWKVFTFCENTLIPVALIDDKILCMHGGSPDLMABDNOR----- 202

Qy 297 PPLTDGEPLDKTENQKQIFDMDWNSDP-QATMGCVPTLRRGAGWFGPDVTDNFQRHRSLY 355  
Db 203 -----PTDPVPGQGLCDLWSDPGKOVQGWGMND-RGVSTTFGADKVSFQLRHDLDL 254

Qy 356 VIRSHECKNGHFRPHMDNKITIFSASNYYAIGSNKGAYTRUNQNLMPHFOYI 415  
Db 255 ICRAHQVVEDGYERFADROLVTTFSAPNYC GefDnAGAMMSUDTLmCSF-QILKPAERK 313

Qy 416 KRUSPKORM 424  
Db 314 XKFMGPNTM 322

RESULT 38

US-10-437-963-111414 ; Sequence 111414, Application US/10437963  
; Publication No. US20040123343A1

; GENERAL INFORMATION:

; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovacic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Wu, Wei  
; APPLICANT: Boukharov, Andrey A.  
; APPLICANT: Barbazuk, Brad  
; APPLICANT: Li, Ping

; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With Plants and Uses Thereof for Plant Improvement

; FILE REFERENCE: 38-21(53221)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; SEQ ID NO: 19487  
; LENGTH: 329

; TYPE: PRT  
; ORGANISM: Glycine max  
; FEATURES:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_17920C.1.pep

US-10-424-599-194787

Query Match 12.4%; Score 427; DB 4; Length 329;  
Best Local Similarity 31.4%; Pred. No. 1. 4e-26;  
Matches 107; Conservative 68; Mismatches 127; Indels 42; Gaps 10;

Qy 105 TPLIDIVPRKKGNGRNRLHPKVALLRRAAKSLKQLENISPVSTAVSQVTQCG 157  
Db 10 INRLKE-RGRPKQV-----LSA-BIKQLCVSRSDFPLRNLLEBAPIKICG 59

Qy 158 DNGKDDLLVLVHNGLPSSSNPYVFNQDFVDRGKRGLEVLILLISLYLAPPNAFLNR 217  
Db 60 DVHQXSYSLRLPEYGFGLPPRN-YLFGLDYVDRGKQSLETICLILAYKIKYPENFLLR 118

Qy 218 GHBDSSPMARGPPIRESKIPRNHLAPIDEYVWPLIGSVLNRLVINGGFS-D 276  
Db 119 GHHECASINRYGPVNCCKRY--NVRWKVFTFCENTLIPVALIDDKILCMHGGSP 175

Qy 277 SPSIDLKISTDKYKVSLRPLTDGEPPLDKTENQKQIFDMDWNSDPRTMGCVPTLRRG 336  
Db 176 LNNLQNQKGLP-----PIEVETGLCDLWSDPSDSDIRGGENEGRV 220

Qy 337 WFGPDVTDNFQRHRSLYVIRSHECKNGHFRPHMDNKITIFSASNYYAIGSNKGAYTR 396  
Db 221 YTFGADKVSFQLRHDLDLICRAHQVVEDGYERFADROLVTTFSAPNYC GefDnAGAMMT 280

Qy 397 LNQNLMPHFOYI 440  
Db 281 VDTELVCSF-QILKPAERK 319

RESULT 40

US-10-425-115-29336 ; Sequence 29336, Application US/10425115

Publication No. US20040314272A1  
 GENERAL INFORMATION:  
 APPLICANT: La Rosa, Thomas J.  
 APPLICANT: Kovalic, David K.  
 APPLICANT: Zhou, Yihua  
 APPLICANT: Cao, Yongwei  
 TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with  
 TITLE OF INVENTION: Plants  
 FILE REFERENCE: 38-21(5322)B  
 CURRENT APPLICATION NUMBER: US /10/425,115  
 CURRENT FILING DATE: 2003-4-28  
 NUMBER OF SEQ ID NOS: 369326  
 SEQ ID NO 29336  
 LENGTH: 322  
 TYPE: PRT  
 ORGANISM: Zea mays  
 FEATURE: OTHER INFORMATION: Clone ID: MRP4577\_30603C.1.pep  
 US-10-425-115-29336

Query Match 12.4%; Score 426.5; DB 4; Length 322;  
 Best Local Similarity 30.1%; Pred. No. 1.5e-26;  
 Matches 111; Conservative 68; Mismatches 135; Indels 55; Gaps 13;

Qy 60 PDAAGRNQYQSAHVSYLDKDLVPERGIVNAK--IEPIRKHIDLIDVPRKG 117  
 Db 5 PAAGG--OGGGCDIAML--DDITIRLLEVRTRAPGKQVQLESBRIOCTVSR--- 54

Qy 118 NRIAHPKTVALLRBAAKSLKLQPNISVUSTAVSQVTCGDLGKQDUDLVLUHKNGLPS 177  
 Db 55 .....AELSQ.....PNLLEAPTK---LGDIHQGQPSDLALUPRYCGFPP 94

Qy 178 SSNPYVNGDFVDRKGKGLVEVLLLSLYLAFFNAVLNRGEHDSDYNNARYGPRREVES 237  
 Db 95 EAN-YLFGLDYDRKGKOSLETICCLLLAYKIKPENFFLRLRGHECASINRYIGFIDECKR 153

Qy 238 KYPRNHRKILAFIDEVTRWLPGSVLMSSRVLTHGGES-DPSFLDKSIRGKVSLR 296  
 Db 154 RF---NTRLWKVFTECENTWLPVALIDDKLICMHGGLSPDIAHLDEIKNOR----- 202

Qy 297 PPLTDGERPLDKIEWQOIFDIMWSDP-QATMGCVPUITRGAGWFGPDVTDLFLQRHLSY 355  
 Db 203 .....PTDVDPQGLCQDLMASDGPQDAQGMOND-RGVSVTGFADKVSFLQKDLD 254

Qy 356 VRSHECKPNQHGHEFMHDNKITITFSASNYYAIGSNKGRAYTRNLNOLOMPHPWQYISAQOT 415  
 Db 255 ICRAHQVEDGIEFFPADRQLVITFSAPNYCIGEFDNAGAMMSVDETLMCSF-OILKPAERK 313

Qy 416 KRLSFKDM 424  
 Db 314 NKPMGSNKM 322

Copyright (c) 1993 - 2006 Compugen Ltd.

### OM protein - Protein search, using SW model

Run on: January 20, 2006, 19:52:48 ; Search time 32 Seconds

Scoring table: BLOSUM62 ; Gapext 0.5

Searched: 70606 seqs, 10133881 residues

Total number of hits satisfying chosen parameters: 70606

Minimum DB seq length: 0

Maximum DB seq length: 0

Perfect score: US-09-463-733-1

Sequence: 1 MDENAIRAALPIQKWMYRRHQ..... VEHDDIDPTDCESKVDPKLS 661

Scoring table: Gapop 10.0 , Gapext 0.5

Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0

Maximum DB seq length: 0

Post-processing: Minimum Match 0% Maximum Match 100%

Database : Listing first 45 summaries

Published Applications AA\_New:\*

```
1: /cgn2_6/ptodata/1/pubpaas/US08_NEW_PUB.PEP:*
2: /cgn2_6/ptodata/1/pubpaas/US06_NEW_PUB.PEP:*
3: /cgn2_6/ptodata/1/pubpaas/US07_NEW_PUB.PEP:*
4: /cgn2_6/ptodata/1/pubpaas/FCI_NEW_PUB.PEP:*
5: /cgn2_6/ptodata/1/pubpaas/US09_NEW_PUB.PEP:*
6: /cgn2_6/ptodata/1/pubpaas/US10_NEW_PUB.PEP:*
7: /cgn2_6/ptodata/1/pubpaas/US11_NEW_PUB.PEP:*
8: /cgn2_6/ptodata/1/pubpaas/US60_NEW_PUB.PEP:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

### SUMMARIES

Result No. Score Query Match Length DB ID Description

Result No.	Score	Query	Match Length	DB ID	Description
1	405.5	11.8	319	7	US-11-109-156-38 Sequence 38, Appl
2	405.5	11.8	323	6	US-10-870-556-184 Sequence 38, Appl
3	405.5	11.8	323	7	US-11-109-156-37 Sequence 37, Appl
4	391.5	11.4	521	7	US-11-109-156-34 Sequence 34, Appl
5	379	11.0	309	7	US-11-109-156-24 Sequence 24, Appl
6	379	11.0	309	7	US-11-109-156-39 Sequence 19, Appl
7	121	3.5	1614	6	US-10-821-234-903 Sequence 903, Appl
8	111	3.2	548	6	US-10-995-561-810 Sequence 810, Appl
9	111	3.2	567	6	US-10-995-561-813 Sequence 813, Appl
10	111	3.2	559	6	US-10-995-561-812 Sequence 812, Appl
11	108	3.1	709	7	US-11-074-176-158 Sequence 11, Appl
12	107.5	3.1	694	7	US-11-202-516-10 Sequence 10, Appl
13	107.5	3.1	752	7	US-11-202-516-8 Sequence 8, Appl
14	105	3.1	5024	6	US-10-793-626-2964 Sequence 2964, Appl
15	104.5	3.0	1070	7	US/11/062 Sequence 4, Appl
16	104.5	3.0	1095	7	US/11/062 Sequence 7, Appl
17	103	3.0	270	6	US-10-118-590-18 Sequence 18, Appl
18	3.0	1765	6	US-10-052-877-140 Sequence 158, Appl	
19	102	3.0	149	6	US-10-821-234-1535 Sequence 1535, Appl
20	102	3.0	149	7	US-11-124-368A-269 Sequence 268, Appl
21	102	3.0	149	7	US-11-124-368A-269 Sequence 269, Appl
22	102	3.0	171	6	US-10-821-234-935 Sequence 935, Appl
23	101.5	2.9	148	7	US-11-065-943-46 Sequence 46, Appl
24	101	2.9	257	6	US-11-118-590-16 Sequence 16, Appl
25	100.5	2.9	1548	7	US-11-108-172-1095 Sequence 1095, Appl

### ALIGNMENTS

```
RESULT 1
US-11-109-156-38
; Sequence 38, Application US/1109156
; Publication No. US2005025014A1
; GENERAL INFORMATION:
; APPLICANT: Toshio Ota
; APPLICANT: Tetsuo Nishikawa
; APPLICANT: Koji Hayashi
; APPLICANT: Kaoru Otsuka
; APPLICANT: Jun-Ichi Yamamoto
; APPLICANT: Shizuko Ishii
; APPLICANT: Tomoyasu Sugiyama
; APPLICANT: Ai Wakamatsu
; APPLICANT: Keiichi Nakai
; APPLICANT: Tetsuji Furukoshi
; APPLICANT: Shin-ichi Furukoshi
; APPLICANT: Jun-Ichi Nezu
; TITLE OF INVENTION: NOVEL GENES ENCODING PROTEIN KINASE/PROTEIN
; FILE REFERENCE: 0501-099002
; CURRENT APPLICATION NUMBER: US/11/109-156
; CURRENT FILING DATE: 2005-04-19
; PRIORITY APPLICATION NUMBER: US/10/060,065
; PRIORITY FILING DATE: 2002-01-29
; PRIORITY APPLICATION NUMBER: PCT/JP00/05061
; PRIORITY FILING DATE: 2000-07-28
; PRIORITY APPLICATION NUMBER: US 60/159,590
; PRIORITY FILING DATE: 1993-10-18
; PRIORITY APPLICATION NUMBER: US 60/183,322
; PRIORITY FILING DATE: 2000-02-17
; PRIORITY APPLICATION NUMBER: JP 11-248036
; PRIORITY FILING DATE: 1993-07-29
; PRIORITY APPLICATION NUMBER: JP 2000-118776
; PRIORITY FILING DATE: 2000-01-11
; PRIORITY APPLICATION NUMBER: JP 2000-183767
; PRIORITY FILING DATE: 2000-05-02
; PRIORITY APPLICATION NUMBER: JP 2000-241899
; PRIORITY FILING DATE: 2000-06-09
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO: 38
; LENGTH: 319
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-11-109-156-38
```

Query Match 11.8%; Score 405.5; DB 7; Length 319;  
 Best Local Similarity 32.4%; Pred. No. 2.7e-24;  
 Matches 105; Conservative 58; Mismatches 126; Indels 35; Gaps 11;

Qy 102 KHNIDLLIDVPRKKGRN-----LHPKYV-ALIREAAKSJKOLPNISPVSTAVSQVT 154  
 Db 2 KUNIDSIIQRLLLEVRSKPKGVQVOLNEBTRGCLIKSRREFLSQL---PILLEAPLK 56

Qy 155 VCGDLGKGLDILWVLUHKNGLPSSSNPYVNGDFVDRGRKGLEVLILLISLYAPPNAVP 214  
 Db 57 ICGDINGQYDILRFLFEGFPPESN-YLFGLDYVDRGKOSLETICLLAYKIKYPENFF 115

Qy 215 LRGNHEDSUMARYGFTREVESKPRHGRILAFIDEVYMLPLGSVNSRVLVHGGF 274  
 Db 116 LLRGNHECASINRYGYFDCKRRY--NIKLMKTFD-CNCCLPAAIVDEKIFCCHGGL 172

Qy 275 S-DSTSLLDKSIDRKYVSLRPLTDGERPLDKTSMQOIDMWDSP-QATMGCVPTNL 332  
 Db 173 SDLQSMEQIRRIMR-----PTDVPDQGLCDLWSDPDKVLGEND- 216

Qy 333 RAGGVWFGPPDVNTFLORHRISYVIRSHBCKNGHSFMHDKLTITPSASNYTAGGSNK 392  
 Db 217 RGVSFTFGAEVVAFLKHDIDLCRAHQVVEDGYSPPAKQQLVTLFSAPNYCGBFDNAG 276

Qy 393 AVTRLNQOLMPHPIVQYISAASQTK 416  
 Db 277 AMMSVDETTIMCSP-QILKPAEKCK 299

RESULT 2  
 US-10-878-556A-184  
 Sequence 184, Application US/10878556A  
 Publication No. US20050266399A1  
 GENERAL INFORMATION:  
 APPLICANT: Hoffmann La Roche Inc.  
 TITLE OF INVENTION: HCV regulated protein expression  
 FILE REFERENCE: 21762  
 CURRENT APPLICATION NUMBER: US/10/878, 556A  
 CURRENT FILING DATE: 2004-06-28  
 NUMBER OF SEQ ID NOS: 199  
 SOFTWARE: PatentIn version 3.1  
 SEQ ID NO: 184  
 LENGTH: 323  
 TYPE: PRT  
 ORGANISM: Homo sapiens  
 PUBLICATION INFORMATION:  
 DATABASE ACCESSION NUMBER: BW\_hum/pp19\_human  
 DATABASE ENTRY DATE: 1994-06-01  
 US-10-878-556A-184

Query Match 11.8%; Score 405.5; DB 6; Length 323;  
 Best Local Similarity 32.4%; Pred. No. 2.8e-24;  
 Matches 105; Conservative 58; Mismatches 126; Indels 35; Gaps 11;

Qy 102 KHNIDLLIDVPRKKGRN-----LHPKYV-ALIREAAKSJKOLPNISPVSTAVSQVT 154  
 Db 6 KUNIDSIIQRLLLEVRSKPKGVQVOLNEBTRGCLIKSRREFLSQL---PILLEAPLK 60

Qy 155 VCGDLGKGLDILWVLUHKNGLPSSSNPYVNGDFVDRGRKGLEVLILLISLYAPPNAVP 214  
 Db 61 ICGDINGQYDILRFLFEGFPPESN-YLFGLDYVDRGKOSLETICLLAYKIKYPENFF 119

Qy 215 LRGNHEDSUMARYGFTREVESKPRHGRILAFIDEVYMLPLGSVNSRVLVHGGF 274  
 Db 120 LLRGNHECASINRYGYFDCKRRY--NIKLMKTFD-CNCCLPAAIVDEKIFCCHGGL 176

Qy 275 S-DSTSLLDKSIDRKYVSLRPLTDGERPLDKTSMQOIDMWDSP-QATMGCVPTNL 332  
 Db 177 SPDLQSMEQIRRIMR-----PTDVPDQGLCDLWSDPDKVLGEND- 220

Qy 333 RAGGVWFGPPDVNTFLORHRISYVIRSHBCKNGHSFMHDKLTITPSASNYTAGGSNK 392  
 Db 221 RGVSFTFGAEVVAFLKHDIDLCRAHQVVEDGYSPPAKQQLVTLFSAPNYCGBFDNAG 280

RESULT 3  
 US-11-109-156-37  
 Sequence 37, Application US/1109156  
 Publication No. US20050250144A1  
 GENERAL INFORMATION:  
 APPLICANT: Toshio Ota  
 APPLICANT: Takao Isogai  
 APPLICANT: Tomoatsu Sugiyama  
 APPLICANT: Koji Hayashi  
 APPLICANT: Kaoru Otsuka  
 APPLICANT: Jun-Ichi Yamamoto  
 APPLICANT: Shizuko Ishii  
 APPLICANT: Ai Wakamatsu  
 APPLICANT: Keiichi Nagai  
 APPLICANT: Tetsuji Otsubo  
 APPLICANT: Shin-ichi Punhashi  
 APPLICANT: Chiaki Senoo  
 APPLICANT: Jun-Ichi Nezu  
 TITLE OF INVENTION: NOVEL GENES ENCODING PROTEIN KINASE/PROTRIN  
 FILE REFERENCE: 06501-099002  
 CURRENT APPLICATION NUMBER: US/11/109, 156  
 PRIOR APPLICATION NUMBER: US/10/060, 065  
 PRIOR FILING DATE: 2002-01-29  
 PRIOR APPLICATION NUMBER: PCT/JP00/05061  
 PRIOR FILING DATE: 2000-07-28  
 PRIOR APPLICATION NUMBER: US 60/159, 590  
 PRIOR FILING DATE: 1999-10-18  
 PRIOR APPLICATION NUMBER: JP 60/183, 322  
 PRIOR FILING DATE: 2000-02-17  
 PRIOR APPLICATION NUMBER: JP 11-248036  
 PRIOR FILING DATE: 1999-07-29  
 PRIOR APPLICATION NUMBER: JP 2000-118776  
 PRIOR FILING DATE: 2000-01-11  
 PRIOR APPLICATION NUMBER: JP 2000-183767  
 PRIOR FILING DATE: 2000-03-02  
 PRIOR APPLICATION NUMBER: JP 2000-241899  
 PRIOR FILING DATE: 2000-06-09  
 NUMBER OF SEQ ID NOS: 43  
 SOFTWARE: PatentIn Ver. 2.0  
 SEQ ID NO: 37  
 LENGTH: 323  
 TYPE: PRT  
 ORGANISM: Homo sapiens  
 US-11-109-156-37

Query Match 11.8%; Score 405.5; DB 7; Length 323;  
 Best Local Similarity 32.4%; Pred. No. 2.8e-24;  
 Matches 105; Conservative 58; Mismatches 126; Indels 35; Gaps 11;

Qy 102 KHNIDLLIDVPRKKGRN-----LHPKYV-ALIREAAKSJKOLPNISPVSTAVSQVT 154  
 Db 6 KUNIDSIIQRLLLEVRSKPKGVQVOLNEBTRGCLIKSRREFLSQL---PILLEAPLK 60

Qy 155 VCGDLGKGLDILWVLUHKNGLPSSSNPYVNGDFVDRGRKGLEVLILLISLYAPPNAVP 214  
 Db 61 ICGDINGQYDILRFLFEGFPPESN-YLFGLDYVDRGKOSLETICLLAYKIKYPENFF 119

Qy 215 LRGNHEDSUMARYGFTREVESKPRHGRILAFIDEVYMLPLGSVNSRVLVHGGF 274  
 Db 120 LLRGNHECASINRYGYFDCKRRY--NIKLMKTFD-CNCCLPAAIVDEKIFCCHGGL 176

Qy 275 S-DSTSLLDKSIDRKYVSLRPLTDGERPLDKTSMQOIDMWDSP-QATMGCVPTNL 332  
 Db 177 SPDLQSMEQIRRIMR-----PTDVPDQGLCDLWSDPDKVLGEND- 220

RESULT 4  
US-11-09-156-34  
Sequence 34, Application US/1109156  
Publication No. US20050250144A1  
GENERAL INFORMATION:  
 APPLICANT: Toshio Ota  
 APPLICANT: Takao Isogai  
 APPLICANT: Tetsuo Nishikawa  
 APPLICANT: Koji Hayashi  
 APPLICANT: Kaoru Otsuba  
 APPLICANT: Jun-Ichi Yamamoto  
 APPLICANT: Shizuko Ishii  
 APPLICANT: Tomoyasu Sugiyama  
 APPLICANT: Ai Wakamatsu  
 APPLICANT: Keiichi Nagai  
 APPLICANT: Tetsuji Otsuki  
 APPLICANT: Shin-Ichi Funahashi  
 APPLICANT: Chiaki Senoo  
 APPLICANT: Jun-Ichi Nezu  
 TITLE OF INVENTION: NOVEL GENES ENCODING PROTEIN KINASE/PROTEIN PHOSPHATASE  
 FILE REFERENCE: 06/01-059002  
 CURRENT APPLICATION NUMBER: US/11/109,156  
 CURRENT FILING DATE: 2005-04-19  
 PRIORITY APPLICATION NUMBER: US/10/060,065  
 PRIORITY FILING DATE: 1999-10-18  
 PRIORITY FILING DATE: 2002-01-29  
 PRIORITY FILING DATE: 2000-02-17  
 PRIORITY FILING DATE: 1999-07-29  
 PRIORITY FILING DATE: 1999-10-18  
 PRIORITY FILING DATE: 1999-07-28  
 PRIORITY FILING DATE: 2000-01-11  
 PRIORITY FILING DATE: 2000-05-02  
 PRIORITY FILING DATE: 2000-06-09  
 PRIORITY FILING DATE: 2000-06-09  
 NUMBER OF SEQ ID NOS: 43  
 SOFTWARE: Patentin Ver. 2.0  
 SEQ ID NO 34  
 LENGTH: 521  
 TYPE: PRT  
 ORGANISM: Homo sapiens  
 US-11-109-156-34

Query Match 11.4%; Score 391.5%; DB 7; Length 521;  
 Best Local Similarity 31.1%; Pred. No. 6.8e-23; Gaps 10;  
 Matches 101; Conservative 58; Mismatches 119; Indels 47;

Qy 102 KMHDLIDIVPRKGNGNLLHPKVVALRRAAKSIKQPNISPVSTAVSQQTVCGLHG 161  
 Db 40 KPRVDILKHLME--GRIBEBSVALRIRITGASILRQENKLDDA---PVTWCGDING 93

Qy 162 KLDLILVVLKNGLPSSSNPYFENGDFVDRKGKQGLDVLVLLSLYLAPNPNAVLRNGH 221  
 Db 94 QFPFLMKLFLFEGSSPANTR-YLFQGDYDGRGSIECIVLYWALKILYPTKLFLRGH 152

Qy 222 DSVMARYCFIRETESKVRNHRKILARPDVTRWLPLPKGSVLMRSVLUVHGGPS-DSPLSL 280  
 Db 153 CRHLLTEYFIFKQBCKIKY--SERVYDACAQFDCLPLAAMNOQFLCVHGGLSPEINTL 209

RESULT 5  
US-11-09-156-24  
Sequence 24, Application US/1109156  
Publication No. US20050250144A1  
GENERAL INFORMATION:  
 APPLICANT: Toshio Ota  
 APPLICANT: Takao Isogai  
 APPLICANT: Tetsuo Nishikawa  
 APPLICANT: Koji Hayashi  
 APPLICANT: Kaoru Otsuba  
 APPLICANT: Jun-Ichi Yamamoto  
 APPLICANT: Shizuko Ishii  
 APPLICANT: Tomoyasu Sugiyama  
 APPLICANT: Ai Wakamatsu  
 APPLICANT: Keiichi Nagai  
 APPLICANT: Tetsuji Otsuki  
 APPLICANT: Shin-Ichi Funahashi  
 APPLICANT: Chiaki Senoo  
 APPLICANT: Jun-Ichi Nezu  
 TITLE OF INVENTION: NOVEL GENES ENCODING PROTEIN KINASE/PROTEIN PHOSPHATASE  
 FILE REFERENCE: 06/01-059002  
 CURRENT APPLICATION NUMBER: US/11/109,156  
 PRIORITY APPLICATION NUMBER: PCT/JP00/05061  
 PRIORITY FILING DATE: 2000-07-28  
 PRIORITY FILING DATE: 2000-01-29  
 PRIORITY FILING DATE: 1999-10-18  
 PRIORITY FILING DATE: 1999-07-28  
 PRIORITY FILING DATE: 2000-02-17  
 PRIORITY FILING DATE: 2000-05-02  
 PRIORITY FILING DATE: 2000-06-09  
 PRIORITY FILING DATE: 2000-06-09  
 PRIORITY FILING DATE: 2000-01-11  
 PRIORITY FILING DATE: 2000-05-02  
 PRIORITY FILING DATE: 2000-06-09  
 PRIORITY FILING DATE: 2000-06-09  
 NUMBER OF SEQ ID NOS: 43  
 SOFTWARE: Patentin Ver. 2.0  
 SEQ ID NO 24  
 LENGTH: 309  
 TYPE: PRT  
 ORGANISM: Homo sapiens  
 US-11-109-156-24

Query Match 11.0%; Score 379; DB 7; Length 309;  
 Best Local Similarity 31.1%; Pred. No. 2.9e-22; Gaps 7;  
 Matches 91; Conservative 62; Mismatches 112; Indels 28; Gaps 7;

Qy 119 RLKPKVVALRREAKSIKQPNISPVSTAVSQQTVCGLHGKLDLILVLLSLYLAPNPNAVLRNGH 178  
 Db 22 QISEQSKVSLCEKAKEILTKESNVQERC---PVTWCGDINGQFHLMELFRIGGSID 77

Qy 179 SNYVPGDFVDRKGKQGLDVLVLLSLYLAPNPNAVLRNGHEDSVNARYCPIREYESK 238  
 Db 78 TN-YLFMDGYDVRGGYSSVETVTLVALVKVYRBRITILRGHESRQITQVGFYDECLRK 136



QY 350 RHLISVVIRSHCK-PNGHEPMDNK---ITIFSAASNYAIGNSKGAYIRNLNQMPH- 404  
Db 1213 HFRKYP--KACNVPEBKEFMKDVKTDLVTV-----SKPVIVIERRISTS 1259  
QY 405 -FVOYTSAASTOKRKLSPKORMGIV----BSALKELAVURRD-----HRBLED 448  
Db 1260 LLEBHDIAPEBKNDLSELLSLGSIGEVTPTVSPLGEAGAVDPPNPKANTLSQLSKTBISL 1319  
QY 449 EF-RKVDPDKDSGYISHWCKMENVTKLGLPWRLLDKLARGTDSQKVNTRTLDT 507  
Db 1320 VLTSKYQJEDGRAIDS---SJMKTQGLID--VIRNQ--PG-----NTLBLET 1364  
QY 508 DVTELEADGMWSMDALYANKSLVATFNIDADSGETDEPETADLW-ARMGAY 566  
Db 1365 PATAQOEVUDHTDM---VSRA-----MDSR-----TPBMKHSQSMEDAOPLEQ 1408  
QY 567 SKAEMLKCRMMDLNG---DEKVD-LNEFRAFRSLDRHKEQDQE----- 608  
Db 1409 KCKKIORNLRLTBLQTHVSSENQDITNEAKDNRQVIRYKRLAKLQQTQNALN 1468  
QY 609 -----NIRRST-----GRPSVAKTADPTVILLADKISK 637  
Db 1469 KGAAYFVERQINQYDVTIKCLONLKRNTSILCKDCKGPKGAKA-KPKYTAALHB 1527  
QY 638 NTLUVBD 645  
Db 1528 KOVLDID 1535

RESULT 8  
US-10-995-561-810  
; Sequence 810, Application US/10995561  
; Publication No. US20050272054A1  
; GENERAL INFORMATION:  
; APPLICANT: CARGILL, Michele et al.  
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH  
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF  
; FILE REFERENCE: C1001559  
; CURRENT APPLICATION NUMBER: US/10/995,561  
; CURRENT FILING DATE: 2004-11-24  
; NUMBER OF SBO ID NOS: 85702  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO: 813  
; LENGTH: 567  
; TYPE: PRX  
; ORGANISM: Homo sapiens  
; US-10-995-561-813  
Query Match 3.2%; Score 111; DB 6; Length 567;  
Best Local Similarity 18.6%; Pred. No. 0.42; Matches 86; Conservative 18.6%; Pred. No. 0.42; Mismatches 175; Indels 134; Gaps 19;  
Matches 86; Conservative 18.6%; Pred. No. 0.42; Mismatches 175; Indels 134; Gaps 19;  
QY 273 GFSDFSTSDLKLISIDRKYVSLRPP---LTGCBPLKTEWQDIFDWTMWDQATMGCVPN 330  
Db 110 GPS-SFRSLVPPER-LHLSLHKSSKRVINGPEKNSSSSPSSVDAASGPRLK----- 161  
; CURRENT FILING DATE: 2004-11-24  
; NUMBER OF SBO ID NOS: 85702  
; LENGTH: 548  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO: 810  
; TYPE: PRX  
; ORGANISM: Homo sapiens  
; US-10-995-561-810  
Query Match 3.2%; Score 111; DB 6; Length 548;  
Best Local Similarity 18.6%; Pred. No. 0.4; Matches 86; Conservative 175; Indels 134; Gaps 19;  
Matches 86; Conservative 18.6%; Pred. No. 0.4; Matches 86; Conservative 18.6%; Pred. No. 0.4; Mismatches 175; Indels 134; Gaps 19;  
QY 273 GFSDFSTSDLKLISIDRKYVSLRPP---LTGCBPLKTEWQDIFDWTMWDQATMGCVPN 330  
Db 91 GPS-SPARSLVPPER-LHLSLHKSSKRVINGPEKNSSSSPSSVDAASGPRLK----- 142  
QY 331 TURGAGWFGPVTDNFLQRHLSYVIRSHCKPNG---HBFMDHNKL----- 376  
Db 143 ---SSGALYGRPRSTPFNSHRQLVSK---APSGDLDKSELSFKNQKLPPTPLKT 196  
QY 377 ---TIFSASNYAIGNSKGAYIRNLNQMPHVOYTSAASTOKRKLSPKORMGIV- 433  
Db 216 EAKSFISQYRITPAKKOFD-----QRIEATQTB-JSPKSBLGTAETKMT 264  
Db 434 EIAVRMEDHR-----DELEDFERY----- 453  
Db 265 DSEMNIKOASNCVTDYAKEKETAPLPLEGHSTWDIEKDHALQHSSPRAMCQSYLSKPPSTR 324  
QY 454 ---DPSDGVISISHWCKMENVTKLGLPWRLLDKLARGTDSQKVNTRTLDT 504  
Db 325 KLYSDBELLYSFIB--DVTEBILKGLFSNRLPLRPERHKHBLERKMRHLHV 382  
QY 505 LDTDVILAEADGMWSMDALVANKSASVAIFIADNSGETLDBEFETADLVLVAMMPG 564  
Db 383 LKVGDGTSEBNSVKQD-----VDMNUVDFERKAGSBRPNELKNESETVTCRQ 433  
QY 565 AVSKAENLKECRMMDLNGDKVDSLNE-----PLAFRSLDRH--KSOODENIRR 613  
Db 434 QVOKA-----LDMILSAPKDNEIPPSPEFFMFIYKSKHSEGVIQOQNDETMLT 485  
QY 614 ST---GRPSVAKTADPTVILLADKISK 652  
Db 486 STUDENRPSISLDTET-----SVNVIEGSDPPEKE 519  
RESULT 10  
US-10-995-561-812  
; Sequence 812, Application US/10995561

Publication No. US20050272054A1  
 GENERAL INFORMATION:  
 APPLICANT: CARGILL, Michele et al.  
 TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH  
 TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF  
 FILE REFERENCE: CL001559  
 CURRENT APPLICATION NUMBER: US/10/995,561  
 CURRENT FILING DATE: 2004-11-24  
 NUMBER OF SEQ ID NOS: 85702  
 SOFTWARE: FastSEQ for Windows Version 4.0  
 SEQ ID NO: 812  
 LENGTH: 599  
 TYPE: PRT  
 ORGANISM: Homo sapiens  
 US-10-995-561-812

Query Match 3.2%; Score 111; DB 6; Length 599;  
 Best Local Similarity 18.6%; Pred. No. 0.46; Matches 86; Conservative 67; Mismatches 175; Indels 134; Gaps 19;

QY 273 GFSDFSTSLDLIKS1DRGKVYSILRPP--LITDGPRLKTEWQIDIMSDPQATMGCVPN 330  
 Db 142 GPS-SPARS-LVWSSER-LHLSHAKSKVITYPEPKNSSSSSVDAASGRKL----- 193  
 QY 331 TLRGAGWFGPDTDNFLQRHLSYVRSHECKPNG----HEFMNDKLT----- 376  
 Db 194 --SSGALYGRPRSTPPNSHARFOLVISK--APSDLDGUSLFSNKQDLPPTPLKT 247  
 QY 377 --TIFASNYNTAIGSNKGAYIRLNQOLMPHFWOYI-SAASTKRSFKORNIGIVSSALK 433  
 Db 248 EAKSFSLSQRYTTPAKKKDFID-----ORIAEATQE-USPKSISLGATKKNMT 296  
 QY 434 ELAVMRHDHR-----DELEBFRKY----- 453  
 Db 297 DSEMNITQASNCVTYDAKEKIALPLPLRGHDSTWDEKTDQAHSSPRAMCQSYSLKPSTR 356  
 QY 454 ---DPKDSGYISISHACKVHENVTKGL---PWRLLRKLAQPGTDSOKVNNTLTL 504  
 Db 357 KYSDBSELLYLISFIE--DVDEILKGLFNRFLERFLERHKQKMLBEEBKMHLHV 414  
 QY 505 LTDVTILEADGMSVNDALYANKASILVAINTIDADNSGETTDLFETADILYAHMPC 564  
 Db 415 LKVDLGTTSERENVKOND-----VDMANVFDFEKAGNSEPNBLNESEVITQBERQ 465  
 QY 565 AYSKAEMLEKRMMDANGDKYDLSN-----EFLAFRISDLHR--REQODENIRR 613  
 Db 466 QOKA-----LDMILSAPKDENSFFPSIPEPFPIYKXHSKBCVITQOVDNSTNLET 517  
 QY 614 ST--GRPSVAKTATPVTLADKISKNTLYVEHDIDPTCCE 652  
 Db 518 STLDENHPSISLSDTBET-----SVNTVIGBOSDPEKVR 551

RESULT 11  
 US-11-074-176-158  
 Sequence 158, Application US/11074176  
 Publication No. US20050250115A1  
 GENERAL INFORMATION:  
 APPLICANT: Klaenhammer, Todd R.  
 APPLICANT: Russell, William M.  
 APPLICANT: Altermann, Eric  
 APPLICANT: McAuliffe, Olivia  
 APPLICANT: Peril, Andrea Azcarate  
 TITLE OF INVENTION: Nucleic Acid Sequences Encoding Proteins and Uses Therefore  
 TITLE OF INVENTION: Stress-Related Proteins and Uses Thereof  
 FILE REFERENCE: 5051-94  
 CURRENT APPLICATION NUMBER: US/11/074,176  
 CURRENT FILING DATE: 2005-03-07  
 PRIOR FILING DATE: 60/551,161  
 NUMBER OF SEQ ID NOS: 381  
 SOFTWARE: FastSEQ for Windows Version 4.0

Query Match 3.1%; Score 108; DB 7; Length 709;  
 Best Local Similarity 17.4%; Pred. No. 1; Matches 113; Conservative 97; Mismatches 176; Indels 264; Gaps 31;

QY 83 DIVERBGDIVAKISLPI-RKHNDLIDYF-RKRGNRHLPKVALLREAKSLKQLP 140  
 Db 68 DLVVB---AKNNKFDEVIGRDEQDINVIETSRKRNPNV----- 104  
 Db 105 LIGPAGVGKTSIVEGLAERTAAGNYPAKOMAMHIIISVNINDMVAGSS--LRGSPERLK 161  
 QY 141 NISPVSTAVSQV-----TVCGDHLGKLDL-LVTLNLQHGPSSNPVYENGDTVRGK 193  
 Db 194 RGLEBVULLLISLYLAPPNAVFLNRGNHEDSVMANARYGFIREVESKYPRNKRILAPIDEV 253  
 Db 162 KVTD----- 253  
 QY 254 YRWLPIGSVANSRVLIVHGGFSDSTSLDLIKS1DRGKVYSILRPLTDSR-----PLDFT 308  
 Db 181 HN-----IVGAGSTDSEN-----NINGDAANILKPALASGBELKLIGATTIS 220  
 QY 309 EWQOQFDIMSDP-----QATMGCVPTLURGAGWFGPDTDNFLQRHLSYVRSBBC 362  
 Db 221 EFORT-----ERKDPLSRRFQVQVPRST-----DVAIKTILGLUCKYSDYHV- 265  
 QY 363 KPGNHFMHDNKKITIFASNYNTAIGSNKGAYIRLNQOLMPHFWOYI-SAQTKL----- 418  
 Db 266 -----KVTDSLKAVELSERVY-----QRY-----LEDKAIDIMDAGAKKLVQP 309  
 QY 419 -----SFKFORMGIVVESAALK-----ELAVMRDHDRDELEBFRKYDDPKDOSGYTISHWC 467  
 Db 310 TDEKSLKNOISALEBKKAERAKAEDYKAAEKEKIASLISQSLKSVDSKNTPEVAKDIY 369  
 QY 468 KVMENNTKULGPWRLRDKLAQPGTISQK-----UNYNRNTLUDTPVLEAS----- 514  
 Db 370 QIBOKTKIPMS-ELHADAEQKNTLDAKLUKRNVNTDORADV1-TDAIAKQTQFDSDR 427  
 QY 515 -----ADGMSVNDALYANKASILVAINTIDADNSGETTDLF--ETA 554  
 Db 428 PTGSELUITGTGPGVKTELAKOLAO-LFNCNKEH1-----RADMSEYQDEMA 473  
 QY 555 IDLVYAHMPC-----AYS-----KA-----EMLBKCRMMDL 580  
 Db 474 VNKLG1GASPGVYGIORGQ1TEKVRHOPSYLILDFEIKANPQVNALLQIMDDRITDA 533  
 QY 581 NG-----DGKYLNEPLA-----FRUSDLHR 602  
 Db 534 QGRTVSSFDUTLIMTSNAGFSKLDLEGKVDQDKLUSALENYFRIPEFLNR 583

RESULT 12  
 US-11-202-516-10  
 Sequence 10, Application US/11202516  
 Publication No. US20060008465A1  
 GENERAL INFORMATION:  
 APPLICANT: Steinal, Lucilla  
 APPLICANT: Mouritsen, Soren  
 APPLICANT: Gautam, Anand  
 APPLICANT: Dalm, Iben  
 APPLICANT: Haaning, Jesper  
 APPLICANT: Leach, Dana  
 APPLICANT: Nielsen, Klaus  
 APPLICANT: Karisson, Gunilla  
 APPLICANT: Rasmussen, Peter  
 TITLE OF INVENTION: Novel Methods for Therapeutic Vaccination  
 FILE REFERENCE: 4614-0107PUS2  
 CURRENT APPLICATION NUMBER: US/11/202,516  
 CURRENT FILING DATE: 2005-08-11

PRIOR APPLICATION NUMBER: US 09/806,703 ;  
 PRIOR FILING DATE: 2001-04-04 ;  
 PCT/DK99/00525 ;  
 TITLE OF INVENTION: Novel Methods for Therapeutic Vaccination  
 FILE REFERENCE: 4514-0107PUS2 ;  
 CURRENT APPLICATION NUMBER: US/11/202.516 ;  
 CURRENT FILING DATE: 2005-08-11 ;  
 PRIORITY NUMBER: US 09/806,703 ;  
 PRIORITY FILING DATE: 2001-04-04 ;  
 PRIORITY NUMBER: PCT/DK99/00525 ;  
 PRIORITY FILING DATE: 1999-10-05 ;  
 PRIORITY APPLICATION NUMBER: DK 1998 01261 ;  
 PRIORITY FILING DATE: 1998-10-05 ;  
 PRIORITY APPLICATION NUMBER: DK 1998 01251 ;  
 PRIORITY FILING DATE: 1998-10-05 ;  
 PRIORITY APPLICATION NUMBER: US 60/105,011 ;  
 PRIORITY FILING DATE: 1998-10-20 ;  
 PRIORITY APPLICATION NUMBER: PCT/DK99/00525 ;  
 PRIORITY FILING DATE: 1999-10-05 ;  
 PRIORITY APPLICATION NUMBER: DK 1998 01251 ;  
 PRIORITY FILING DATE: 1998-10-05 ;  
 PRIORITY APPLICATION NUMBER: US 60/105,011 ;  
 SEQ ID NO: 10 ;  
 LENGTH: 694 ;  
 TYPE: PRT ;  
 ORGANISM: Mus musculus ;  
 US-11-202-516-10 ;  
 Query Match 3.1%; Score 107.5%; DB 7; Length 694;  
 Best Local Similarity 19.9%; Pred. No. 1.1.; Mismatches 194; Indels 173; Gaps 29;  
 Matches 109; Conservative 72; MisMatches 194; Length: 694 ;  
 Qy 38 EYASBQDQARLYKEFPNDLKHMPQAGRKQYQSAHWSVLDKDOLVERG-DIVNAKI 96 ;  
 Db 4 BEFLHELAENIKKELYNPR-TPHLAGTQNPELAQOI-----HDQMKERFGDLV--- 52 ;  
 Qy 97 ELPIRKNHHTDLIDVFRKCKGNRLRHPKVALLIREAKK---SLKQLP---NIS---P 144 ;  
 Db 53 EL---SHYDVILSY----PNKTHPNNTISINEDGNEIFKTSLSBOPPPGYENISDVP 103 ;  
 Qy 145 VSTAVSQQTIVCGDL---HGKLDDLJWLHKNGLPPSSNPVYFNGDPFDVDRGGLVLL 200 ;  
 Db 104 PYSAFSPQGTPEGDLVYVNYARTEDPFKLERMKISCSGKIVIARYGVFRGN----- 156 ;  
 Qy 201 LILSILYLAFFNAVFLNRGNHEDSUMMARYGFIREVESEKIPRNHRKILAFDEVYRM-LPL 259 ;  
 Db 157 MYKNAQLAGAKGML---YSPP---ADY-FPVAVKS-YPDG-----WNLPG 194 ;  
 Qy 260 GSVINRSRVTLVHGGPSDSTSBLDIKSIDRGKVTSLRPLTDGPBLKTEWOO----- 312 ;  
 Db 195 GGVORGNVNLNGLA-----GDPLTGPY PANBHAYRHELTNAVG 232 ;  
 Qy 313 -----IDIMMWDPOATMGCVPTLIRGAGW-----FGDPVDTDFLQRERLSTVIR 358 ;  
 Db 233 LPSTVHPVPIGDAQKLLHEHMGGPAPPDSSWKGLKVPYNGPGFACNP-----S 282 ;  
 Qy 359 SHECKPNCHEFMDNKITIFSASNYYAIGSKGAYIRLNQLOMPHFQYISAASQTKRL 418 ;  
 Db 283 TOKYOMIHSYKVTRI-----YVIGITLGKA-----LEPD--RYVILGHRDWA 325 ;  
 Qy 419 SF----KORMGIVVESALKELAVERMDRDELEDFRKYDKDSCGYSYISISHCKMENV 473 ;  
 Db 326 VFGCIDPOSGAAVHVETVRSFGTLKKKGRRPRTILFASWDAEFGLGSTEWABBS-- 383 ;  
 Qy 474 TKUGLPWRRL-----DKLAPGTSQKVN-----YNRTLDDLTDTVLAREA 515 ;  
 Db 384 -----RLLQERGVAYINADSSIEGNYTIRVDCPTPLMSLVNLTKEIQSPDEGPF-- 433 ;  
 Qy 516 DGMSVMDA 523' ;  
 Db 434 -GKSLYDS 440 ;  
 RESULT 13 ;  
 US-11-202-516-8 ;  
 Sequence 8, Application US/11202516 ;  
 Publication No. US2006008465A1 ;  
 GENERAL INFORMATION:  
 APPLICANT: Steinaa, Lucilla ;  
 APPLICANT: Mouritsen, Soren ;  
 APPLICANT: Gautam, Anand ;  
 APPLICANT: Dalum, Iben ;  
 APPLICANT: Haaning, Jesper ;  
 APPLICANT: Leech, Dana ;  
 APPLICANT: Nielsen, Klaus ;  
 ;  
 RESULT 14 ;  
 US-10-793-626-2964 ;  
 Sequence 2964, Application US/10793626 ;  
 Publication No. US20050255478A1 ;  
 GENERAL INFORMATION:  
 APPLICANT: KIMMERLY, WILLIAM JOHN ;



Db 866 FNTELSQSNEER---YKIQSYS---BYLKDFWMGNPLMVMKYYMFN 906  
 Qy 441 DHRDELEDER-RKYDPKDSCYISTSHWCKVMENTVKLGLPWRLLDK-----LAP 489  
 Db 930 SNSQSINDDIVRKED---VIVIDPF----NLNOBWRVVTYKYPKEBEKLFLAP 976  
 Qy 490 GTDSOKVNRTTLDLTD 508  
 Db 977 ISDSDFP-YN-TIQKEYD 993

RESULT 16  
 US/11/062  
 Sequence 7, Application US/11062471A  
 Publication No. US20050255093A1  
 GENERAL INFORMATION:  
 APPLICANT: SHONE, Clifford Charles  
 APPLICANT: SUTTON, John Mark  
 APPLICANT: HALLIS, Bassem  
 APPLICANT: SILMAN, Nigel  
 TITLE OF INVENTION: Delivery of Superoxide Dismutase to Neuronal Cells  
 FILE REFERENCE: 1581\_080001  
 CURRENT APPLICATION NUMBER: US/11/062,471A  
 CURRENT FILING DATE: 2005-02-22  
 PRIOR APPLICATION NUMBER: 09/831,050  
 PRIOR FILING DATE: 1999-11-05  
 PRIOR APPLICATION NUMBER: PCT/CB99/03699  
 PRIOR FILING DATE: 1999-11-05  
 PRIOR FILING DATE: 1998-11-05  
 NUMBER OF SEQ ID NOS: 11  
 SOFTWARE: PatentIn Ver. 2.1  
 SEQ ID NO: 7  
 LENGTH: 1095  
 TYPE: PRT  
 ORGANISM: Artificial Sequence  
 OTHER INFORMATION: Construct comprising a mitochondrial leader sequence from Human M  
 US/11/062,471A-7

Query Match 3.0%; Score 104.5; DB 7; Length 1095;  
 Best Local Similarity 18.7%; Pred. No. 3.6.;  
 Matches 115; Conservative 98; Mismatches 182; Indels 213; Gaps 33;

Qy 10 IPIOKNY---RHOARRNRNRRNWNQITONL-----YASQDOPABLYKFND-- 54  
 Db 503 LIVACOWLSTNTQFTIKEGYKALNYQA-QALBIEIKYKINYSEKEKNINIDFNDIN 561

Qy 55 -LIKMPQAGKRNQYQSSAHVSLDDK--DDUVERGDIWAKIELPRKNDLID 110  
 Db 562 SKLNREGINODIDNTNPNFGCSVSYIMKNIPLATEKLFDNT----LKONLNYIDE 616

Qy 111 VFRKGGRGNRHLPKYVALLBAAKSILQJENISPVASTAVSQVTWCGDLIGKL--- 163  
 Db 617 -NKY---LIGSAETEKSINKVLTIMPDLSTYNTDILLEMHNKNSBILAN 667

Qy 164 -----DPLIVWAKNG-----LSSSNP-----YVFNDFP 188  
 Db 668 ILINRYKOMNLIDLSGYGAKEVVDGVELANDKNOFKLTSANSIKRVTQNQNLIFNSVF 727

Qy 189 VDRGRGLEVILLIJSLYIAFP---NAVELARGHEDSINA----- 227  
 Db 728 LD-----FSVSPWIRPKYKNGDTONYHNEYTINCMKMNKGWKRISTGRNRIW 777

Qy 228 -----RYGFIREVESEKIPRONKHLAPIDEYRW---LPLGSSVLNSRVLVH 271

Db 778 TLIDINGKTISVESEN-IRHDSEY-----INRWEPTITNLANKAVIN 823

Qy 272 GGFSDBSTSL-DLKISDRGKVSIARPLTDGEPLKTE---WQQTDFDIMMSDPORTMGCV 328  
 Db 824 GKLESNTDIOIREVIANGSIIPKL---DGD-IDRTQPIWMKVFSI----- 865

Qy 329 PNTLGAGGVFGDPDVDFQRHRLS-SYVHSHECKPNIGHFMHD-----NKLITFS 380

RESULT 17  
 US-10-118-590-18  
 Sequence 18, Application US/10118590  
 Publication No. US20050277761A1  
 GENERAL INFORMATION:  
 APPLICANT: KENNETH RHODES, MARIA BETTY, HUAI-PING LING, AND FRANK AN  
 FILE REFERENCE: MNI-070  
 CURRENT APPLICATION NUMBER: US/10/118,590  
 CURRENT FILING DATE: 2002-04-08  
 PRIOR APPLICATION NUMBER: US/09/298,731  
 PRIOR FILING DATE: 1999-04-23  
 NUMBER OF SEQ ID NOS: 47  
 SOFTWARE: PatentIn Ver. 2.0  
 SEQ ID NO: 18  
 LENGTH: 270  
 TYPE: PRT  
 ORGANISM: Mus musculus  
 US-10-118-590-18

Query Match 3.0%; Score 103; DB 6; Length 270;  
 Best Local Similarity 25.7%; Pred. No. 0.59; Mismatches 58; Indels 88; Gaps 13; Matches 59; Conservative 25;

Qy 439 MDRDR-----DELBEDPFRKYDPKDSGYVISISHWCKVMENTVKLGLPWRLLDKLA---- 488  
 Db 64 LRPERPRPRLPDPSVEDR-----BLSTVCHRPGLQBLQKTRBLQVLR 112

Qy 489 -----PGTDSOKVNRT-----DLLDTP---VILRAEDENS 519  
 Db 113 GFKNKBCPSGTUNSENFKQTYSQFPQGDSS--NWTATLFNAFDNHDSGSVSPDFVAGLS 170

Qy 520 VM-----DALYAKASLVAIFNVIDANGSGTLEDEBTAINL--VARMGAYSKEM 571  
 Db 171 VILRGTDIQL--NWA---FNUYDLMKDGTCTKEM--LDMKSYDMMCKYTYPAL 220

Qy 572 DEKC-----RMDMDGKVDNLNFLEATPLSLARKBQDENTRR 612  
 Db 221 REAPPRAEVESPFQMDRNGKSVTIESPIESC-----QDENMR 261

RESULT 18  
 US-10-035-877-140  
 Sequence 140, Application US/10055877  
 Publication No. US2005028241A1  
 GENERAL INFORMATION:  
 APPLICANT: Decristofaro, Marc  
 APPLICANT: Padigaru, Muralidhara  
 APPLICANT: Miller, Charles  
 APPLICANT: Tchernev, Velizar  
 APPLICANT: Zheng, Mei  
 APPLICANT: Anderson, David  
 APPLICANT: Ballinger, Robert  
 APPLICANT: Gerlach, Valerie  
 APPLICANT: Spytek, Kimberly  
 APPLICANT: Ratelli, Luca  
 APPLICANT: Kefuda, Ramesh  
 APPLICANT: Guo, Xiaojia

APPLICANT: Zarhusen, Bryan  
 APPLICANT: Andrew, David  
 APPLICANT: Mezes, Peter  
 APPLICANT: Paturjan, Meera  
 APPLICANT: Burgess, Catherine  
 APPLICANT: Eisner, Andrew  
 APPLICANT: Woleenc, Adam  
 APPLICANT: Baumgartner, Jason  
 APPLICANT: Shlimkets, Richard  
 APPLICANT: Gusev, Vladimir  
 APPLICANT: Vernet, Corine  
 APPLICANT: Traupier Jr., Raymond  
 APPLICANT: Pena, Carol  
 APPLICANT: Shenoy, Surash  
 APPLICANT: Li, Li  
 APPLICANT: Caeman, Stacie  
 APPLICANT: Boldog, Ference  
 TITLE OF INVENTION: Novel Polypeptides and Nucleic Acids Encoded Thereby  
 FILE REFERENCE: 21402-251  
 CURRENT APPLICATION NUMBER: US/10/055,877  
 CURRENT FILING DATE: 2002-01-22  
 PRIOR APPLICATION NUMBER: 60/262,892  
 PRIOR FILING DATE: 2001-01-19  
 PRIOR APPLICATION NUMBER: 60/263,598  
 PRIOR FILING DATE: 2001-01-23  
 PRIOR APPLICATION NUMBER: 60/263,799  
 PRIOR FILING DATE: 2001-01-24  
 PRIOR APPLICATION NUMBER: 60/264,117  
 PRIOR FILING DATE: 2001-01-25  
 PRIOR APPLICATION NUMBER: 60/264,139  
 PRIOR FILING DATE: 2001-01-25  
 PRIOR APPLICATION NUMBER: 60/264,478  
 PRIOR FILING DATE: 2001-01-26  
 PRIOR APPLICATION NUMBER: 60/263,351  
 PRIOR FILING DATE: 2001-01-30  
 PRIOR APPLICATION NUMBER: 60/272,870  
 PRIOR FILING DATE: 2001-03-02  
 PRIOR APPLICATION NUMBER: 60/275,990  
 PRIOR FILING DATE: 2001-03-14  
 PRIOR APPLICATION NUMBER: 60/275,927  
 PRIOR FILING DATE: 2001-03-14  
 Remaining Prior Application data removed - See File Wrapper or PALM.  
 NUMBER OF SEQ ID NOS: 512  
 SOFTWARE: PatentIn Ver. 2.1  
 SEQ ID NO 140  
 LENGTH: 1765  
 ; TYPE: PRT  
 ; ORGANISM: Mus musculus  
 US-10-055-877-140

Query Match 3.0%; Score 103; DB 6; Length 1765;  
 Best Local Similarity 21.1%; Pred. No. 9.4; Matches 107; Conservative 54; Mismatches 177; Indels 168; Gaps 26;

Qy 37 LEVSEODAABLYKEFDLKKHMPDAERKKNQYQGSATVSVLDDKDVLEEF--GDIVN 93  
 Db 453 LHMARSQCRKVRY--LVDGAQVEAKADODPLHISARLGKADIVQQLQGASPQ 509

Qy 94 A-----KELPIRKNHDLIDVFRKRGKRNHLPPKVALLRE--AAKSUKQLPNTSP 144  
 Db 510 AAT"SGYTPHLAAREGHED-----VAFLIDHGASLSITTKKGFTP 551

Qy 145 VSTAVSQQTVCGLDKLHKGLPSSNSNPVFGDFVDRQKGL-----196  
 Db 552 LHVAK-----YKLEVASLILQKSASPDAA-----GKSGUTPLHVAH 590

Qy 197 ---EVLLILSLYLAFFPAVPL-----RGHNEDSVM-ARGF-----IR 233  
 Db 591 YDNOKVALILD-OGASPHAAKNGYTPHLIAKRNQMDIAATSLLEYADANAVTROGIA 649

Qy 234 EYESKYPRHRKRLAFIDPYRMPLGSVINSRVLVH-GGFSDSTSLSLUDLSIKSDKTV 292  
 Db 650 SVHLAAQEGH-----VDIV-----SILLSRNVNLNSLNKSGLTPHLAAQBDRVNA 696

Qy 293 STIRRPPTDGEPLDKTEMQQIFDIMMSDPOATMGCVNTLRAGWGRGPDTDNFLQRHR 352  
 Db 697 EVL---WNQGAV-----DAQTKGTYP---LHVQCHYGNIKUVNFIQHS 736  
 Qy 353 LSTVIR-----HRCKPNQSHFMRHMNDKNTTFSNSYYAIGSNKGAYI--RLNQOLMP 403  
 Db 737 AKUNAKTYKNGYTAHQAAQGQTHI-LNVLQNNASBNELTYNTALAIARRIG---- 790  
 Qy 404 HFWQYISAS-----QTKRLSKORMKGIVVESALKELAVRHEHDHOBLEDERRY 453  
 Db 791 ---YISVWDTLKVTBIMTTTTRHKKOMVPET-----MNEVLOMSDDYTRKA 837  
 Qy 454 DP---KDSGYTISH-----WCKV 469  
 Db 838 SAPEKLSKJGEYVSDGERGDKCWTWPKI 863

RESULT 19  
 US-10-821-234-1535  
 ; Sequence 1535, Application US/10821234  
 ; Publication No. US20050255114A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Labat, Ivan  
 ; APPLICANT: stach-Crain, Birgit  
 ; APPLICANT: Andarmani, Susan  
 ; APPLICANT: Tang, Y. Tom  
 ; TITLE OF INVENTION: Methods for Diagnosis and Treatment of Preeclampsia  
 ; FILE REFERENCE: 821A  
 ; CURRENT APPLICATION NUMBER: US/10/821,234  
 ; CURRENT FILING DATE: 2004-04-07  
 ; PRIOR APPLICATION NUMBER: US 60/462,047  
 ; PRIOR FILING DATE: 2003-04-07  
 ; NUMBER OF SEQ ID NOS: 1704  
 ; SOFTWARE: PT-SEQ\_Genes Version 1.0  
 ; SEQ ID NO 1535  
 ; LENGTH: 149  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 US-10-821-234-1535

Query Match 3.0%; Score 102; DB 6; Length 149;  
 Best Local Similarity 28.1%; Pred. No. 0.3; Matches 39; Conservative 20; Mismatches 58; Indels 22; Gaps 6;

Qy 483 LADKLKAGTDSKVNRYTLDLTDWILEBADGM-----SYMDALYAN-KASLVA 533  
 Db 1 MADQI---TEBQIAEFFRAFSIFDKD-----GDTGTTKELGTVMRSLGQPTEARLQD 51

Qy 534 IFNIIDADNSGRTDPERATIDLVHMPGAYSKAEMLEKCRMMDINGDKUDLNEPLE 593  
 Db 52 MINTEVDAGNGTIDPFPLT---NMWAKKMDTSEBEBIREAFRVFDQDNGYVISAELRH 108

Qy 594 AF-RLSDLHRKEQODENIR 611  
 Db 109 VMTNLGSKLTDBEDEVEMIR 127

RESULT 20  
 US-11-124-368A-268  
 ; Sequence 268, Application US/11-124-368A  
 ; Publication No. US20050287559A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Michele Cargill  
 ; APPLICANT: James J. Devlin  
 ; APPLICANT: May Luke  
 ; TITLE OF INVENTION: Genetic Polymorphisms Associated with Vascular Disease, Methods of Detection and Uses Thereof  
 ; TITLE OF INVENTION: Vascular Disease, Methods of Detection and Uses Thereof  
 ; FILE REFERENCE: C1-001524  
 ; CURRENT APPLICATION NUMBER: US/11-124-368A  
 ; CURRENT FILING DATE: 2005-05-09  
 ; PRIOR APPLICATION NUMBER: US 60/568,845  
 ; PRIOR FILING DATE: 2004-05-07

PRIOR APPLICATION NUMBER: US 60/625,936 ; GENERAL INFORMATION:  
; PRIORITY FILING DATE: 2004-11-09 ; APPLICANT: Labat, Ivan  
; NUMBER OF SEQ ID NOS: 21112 ; APPLICANT: Stache-Crain, Birgit  
; SOFTWARE: FastSEQ for Windows Version 4.0 ; APPLICANT: Andarmani, Susan  
; SEQ ID NO: 268 ; APPLICANT: Tang, Y. Tom  
; LENGTH: 149 ; TITLE OF INVENTION: Methods for Diagnosis and Treatment of Preeclampsia  
; ORGANISM: Homo sapiens ; FILE REFERENCE: 821A  
; TYPE: PRT ; CURRENT APPLICATION NUMBER: US/10/821,234  
; US-11-124-368A-268 ; CURRENT FILING DATE: 2004-04-07  
; Best Local Similarity 3.0%; Score 102; DB 7; Length 149; ; PRIOR APPLICATION NUMBER: US 60/462,047  
; Matches 39; Conservative 20; Mismatches 58; Indels 22; Gaps 6; ; PRIOR FILING DATE: 2003-04-07  
; QY 483 LRDKLAPGTSOKUNYNTLTLIDLTDVILVLAEDGM-----SYNDALYAN-KASLVA 533 ; NUMBER OF SEQ ID NOS: 1704  
; Db 1 MADQL--TEEQIAEKFKAFLPKSLFDKD----GDGTITKELGTVMSLQNPTEABQD 51 ; SOFTWARE: pt\_Seq\_Genes Version 1.0  
; QY 534 IFNITDADNSGETLDEPETAIDLIVAHMGAWSKAEMLEKCRMMDLNGDKVLDNLNEPLE 593 ; SEQ ID NO: 935  
; Db 52 MINEVDADGNGTDPPEFT--MMARKMDKTDSBEEIRAFRVPFDKGNGYISAABLH 108 ; LENGTH: 171  
; QY 594 AP-RLSDLHKKRQEODENIR 611 ; TYPE: PRT  
; Db 109 VMTNLGEKLTDEEVDEMIR 127 ; ORGANISM: Homo sapiens  
; RESULT 21 ; US-10-821-234-935  
; US-11-124-368A-269 ; Query Match 3.0%; Score 102; DB 6; Length 171;  
; Publication No. US20050287559A1 ; Best Local Similarity 28.1%; Pred. No. 0, 36; Mismatches 58; Indels 22; Gaps 6;  
; GENERAL INFORMATION: ; Matches 39; Conservative 20; Mismatches 58; Indels 22; Gaps 6;  
; APPLICANT: Michele Cargill ; QY 483 LRDKLAPGTSOKUNYNTLTLIDLTDVILVLAEDGM-----SYNDALYAN-KASLVA 533  
; APPLICANT: James J. Devlin ; Db 23 MADQL--TEEQIAEKFKAFLPKSLFDKD----GDGTITKELGTVMSLQNPTEABQD 73  
; TITLE OF INVENTION: Genetic Polymorphisms Associated with ; QY 534 IFNITDADNSGETLDEPETAIDLIVAHMGAWSKAEMLEKCRMMDLNGDKVLDNLNEPLE 593  
; TITLE OF INVENTION: Vascular Diseases, Methods of Detection and Uses Thereof ; Db 74 MINEVDADGNGTDPPEFT--MMARKMDKTDSBEEIRAFRVPFDKGNGYISAABLH 130  
; FILE REFERENCE: CL001524 ; QY 594 AP-RLSDLHKKRQEODENIR 611  
; CURRENT APPLICATION NUMBER: US/11/124,368A ; Db 131 VMTNLGEKLTDEEVDEMIR 149  
; CURRENT FILING DATE: 2005-05-09 ;  
; PRIOR APPLICATION NUMBER: US 60/568,845 ;  
; PRIOR FILING DATE: 2004-05-07 ;  
; PRIOR APPLICATION NUMBER: US 60/625,936 ;  
; PRIOR FILING DATE: 2004-11-09 ;  
; NUMBER OF SEQ ID NOS: 21112 ;  
; SOFTWARE: FastSEQ for Windows Version 4.0 ;  
; SEQ ID NO: 269 ;  
; LENGTH: 149 ;  
; TYPE: PRT ;  
; ORGANISM: Homo sapiens  
; US-11-124-368A-269 ;  
; Query Match 3.0%; Score 102; DB 7; Length 149;  
; Best Local Similarity 28.1%; Pred. No. 0, 3; Mismatches 58; Indels 22; Gaps 6; ;  
; Matches 39; Conservative 20; Mismatches 58; Indels 22; Gaps 6; ;  
; QY 483 LRDKLAPGTSOKUNYNTLTLIDLTDVILVLAEDGM-----SYNDALYAN-KASLVA 533 ;  
; Db 1 MADQL--TEEQIAEKFKAFLPKSLFDKD----GDGTITKELGTVMSLQNPTEABQD 51 ;  
; QY 534 IFNITDADNSGETLDEPETAIDLIVAHMGAWSKAEMLEKCRMMDLNGDKVLDNLNEPLE 593 ;  
; Db 52 MINEVDADGNGTDPPEFT--MMARKMDKTDSBEEIRAFRVPFDKGNGYISAABLH 108 ;  
; QY 594 AP-RLSDLHKKRQEODENIR 611 ;  
; Db 109 VMTNLGEKLTDEEVDEMIR 127 ;  
; RESULT 22 ;  
; US-10-821-234-935 ; Query Match 2.9%; Score 101.5; DB 7; Length 148;  
; Sequence 935, Application US/10821234 ; Best Local Similarity 28.2%; Pred. No. 0, 32; Mismatches 57; Indels 19; Gaps 5;  
; Publication No. US20050255114A1 ; Matches 37; Conservative 18; Mismatches 57; Indels 19; Gaps 5;  
; QY 491 TPSOKUNYNTLTLIDLTDVILVLAEDGM-----SYNDALYAN-KASLVAIFNITDAD 541 ;  
; Db 5 TBEQIAEKFKAFLPKSLFDKD----GDGTITKELGTVMSLQNPTEABQD 58 ;  
; QY 542 NSCRITDPERMAIDLIVAHMGAWSKAEMLEKCRMMDLNGDKVLDNLNEPLE-RISVL 600 ;  
; Db 59 GNCTTDPPEFT--MMARKMDKTDSBEEIRAFRVPFDKGNGYISAABLRTWMTNGEK 115 ;  
; QY 601 HRKEQODENIR 611 ;

DB 116 LIDDEBDEMIR 126

RESULT 24  
US-10-118-590-16  
; Sequence 16, Application US/10118590  
; GENERAL INFORMATION:  
; APPLICANT: KENNETH RHODES, MARIA BETTY, HUAI-PING LING, AND FRANK AN  
; TITLE OF INVENTION: POTASSIUM CHANNEL INTERACTORS AND USES THEREFOR  
; CURRENT APPLICATION NUMBER: US/10/118,590  
; CURRENT FILING DATE: 2002-04-08  
; PRIOR APPLICATION NUMBER: US/09/298,731  
; PRIORITY FILING DATE: 1999-04-23  
; NUMBER OF SEQ ID NOS: 47  
; SEQ ID NO 16  
; LENGTH: 257  
; TYPE: PRT  
; ORGANISM: Rattus sp.  
; US-10-118-590-16

Query Match 2.9%; Score 101; DB 6; Length 257;  
Best Local Similarity 24.9%; Pred. No. 0; 79; Mismatches 64; Indels 78; Gaps 12;  
Matches 56; Conservative 27; Mismatches 64; Indels 78; Gaps 12;

Qy 439 MRDRH-----DEBUDSERKYPDKPDGYVISISHCKMENVTKLGPWRLRDKLA---- 488  
Db 51 LAPHPRPRLDPDSVDED-----ELSVTCRHEPEGLQLOSGQTKPTEQLQYR 99

Qy 489 -----PGTDSQKUNYNTL-DLTD-----VILRAADGMS 519  
Db 100 GFKNECPSGIVNBENFKQIYSSOPPOGDSS--NYATVFLNAFDTNHDGSVSFEDTAGLS 157

Qy 520 V-MDALYANKSLVATPNIADNSGETLDERETAILL--VAMPGAYSAEMBLKC- 575  
Db 158 VILRGCTDDRLSVA-FNLYDINKGCTIKSEM--LDIMKSIYDMKRYTPALRKEAP 212

Qy 576 -----RMMDLANGSKVLDNEBLFAFRSLDHLRKQSQDENIRR 612  
Db 213 REHVESTPQMDRNKGTVTIEFIREC-----QDENINMR 248

RESULT 25  
US-11-108-172-1095  
; Sequence 1095, Application US/11108172  
; Publication No. US20050260177A1

GENERAL INFORMATION:  
; APPLICANT: Xu, Jiangchun  
; APPLICANT: Lodes, Michael J.  
; APPLICANT: Secrett, Heather  
; APPLICANT: Benson, Darin R.  
; APPLICANT: Meagher, Madeleine Joy  
; APPLICANT: Stolk, John A.  
; APPLICANT: Wang, Tongcong  
; APPLICANT: Jiang, Yuqiu  
; APPLICANT: Smith, Carole L.  
; APPLICANT: King, Gordon E.  
; APPLICANT: Wang, Ajun  
; APPLICANT: Clasper, Jonathan D.  
; APPLICANT: Skeiky, Yasir A. W.  
; APPLICANT: Fanger, Gary R.  
; APPLICANT: Vedrick, Thomas S.  
; APPLICANT: Carter, Darwick  
; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND DIAGNOSIS  
; TITLE OR INVENTION: OF COLON CANCER AND METHODS FOR THEIR USE  
; FILE REFERENCE: 21.0121.471C15  
; CURRENT APPLICATION NUMBER: US/11/108,172  
; PRIORITY FILING DATE: 2005-04-15  
; PRIOR APPLICATION NUMBER: US 10/025,380  
; PRIORITY FILING DATE: 2001-12-19

Query Match 2.9%; Score 100.5; DB 7; Length 1548;  
Best Local Similarity 20.7%; Pred. No. 12; Mismatches 11; Conservative 65; Mismatches 182; Indels 177; Gaps 27;  
Matches 111; Conservative 65; Mismatches 182; Indels 177; Gaps 27;

Qy 138 QPNISPVSTASQSVTQGDIHQKDLVVLVHKGOLPSSNPYVNGDFTRGKGL 197  
Db 77 QPNPRLSNAAATGIA-----GLPSLHNRTVLGPF-----GYH 111

Qy 198 VILLLSUYLATPNAVLN-----RGHEDSVNARYOFIREWEVKYPRNH 243  
Db 112 VSDDVSVETPCCPARTANIRPPGDPVDPDQRGDWVLPFRSRW---DDETRGSPSN 167

Qy 244 KILLAFIDRVVWPLPLGSVLSNLVLTIGFGFDSSTDLSIDRKVSVIRP--LTD 301  
Db 168 PRDLA--NOVTCWLD-GSAIY-----GSHSWS-DALRSRSGGOLASGPDPAFRDS 215

Qy 302 GPRLDKTEWQOIPDIMSDDPQITMGCVPTNLG-----AGVMP--- 339  
Db 216 QPPL-----LMAAPDPATG--QNGPRLYAFGABRNREPFLQLQGLLMFRYN 263

Qy 340 -----GPDVTDNFQLQRHLSYVTRS-----HECKPN-----GHEFMHDNK 374  
Db 264 LMAQLARHPDWEDESFQHAKRKVATYQOMIAVWLPSPLOKTLPEYTYRPFLDPS 323

Qy 375 ITITF-SASNNTAIGSNKGAYTRILNOLMPFIVQYS--MASOTKRLS---EKORMG 425  
Db 324 ISPERFVVVASBOPFSTMVPPGVYMR--NASCHPKVNLKGFOSSQALRVCNNTWIRENP 380

Qy 426 IVESSALKGLAVRMQRDRDLED----EPRKDPDKSGYVISISHCKMENVTKLGP- 479  
Db 381 LMSTQETNELLGMAQSQISELENDIVTDLRQWPGCKFSITDVIASSIQGRQDMELPS 440

Qy 480 -----WRLRDLGLP3DSOKN-----YNRFLDLPDTV--ILEARADG 517  
Db 441 YQALLAFLGDLTPRNW---SDLNPNVDPQVILEATAALYNQDLSQURLLGGCLESHGDP 496

Qy 518 MSYMDALYANKSLVATPNIAD-----NSGETLDEPS---TAIDLVA 560  
Db 497 GLFSATVLDOD-----FVRLRGDRYFTENTINGLSSKKELTDIRNTLDRVILVA 546

RESULT 26  
US-10-467-657-2502  
; Sequence 2502, Application US/10467657  
; Publication No. US20050260581A1  
; GENERAL INFORMATION:

APPLICANT: CHIRON SpA ; TYPE: PRT  
; APPLICANT: FONTANA Maria Rita ; ORGANISM: Lactobacillus acidophilus  
; APPLICANT: PIZZA Mariagrazia ; US-11-074-176-314  
; APPLICANT: MASIGNANI Vega  
; APPLICANT: MONACI Elisabetta  
; TITLE OF INVENTION: GONOCCOCAL PROTEINS AND NUCLEAR ACIDS  
; FILE REFERENCE:  
; CURRENT APPLICATION NUMBER: US/10/467,657  
; CURRENT FILING DATE: 2003-08-11  
; PRIOR APPLICATION NUMBER: GB-0103424.8  
; PRIOR FILING DATE: 2001-02-12  
; NUMBER OF SEQ ID NOS: 9218  
; SOFTWARE: SeqWin99, version 1.04  
; SEQ ID NO: 2502  
; LENGTH: 276  
; TYPE: PRT  
; ORGANISM: Neisseria gonorrhoeae  
; US-10-467-657-2502

Query Match 2.9%; Score 99; DB 6; Length 276;  
Best Local Similarity 18.8%; Pred. No. 1.3; Matches 71; Conservative 51; Mismatches 123; Indels 132; Gaps 17;

Qy 157 GDLHKLDDLWVHKNGLSSSNPVPYRNDPVRGKRGELVLLSLYLAPRNALN 216  
Db 7 GDIQICPDELTTALGKGIGENHGDTLWLQDIDVNRGPKSILT---LDPCCRHENSVQIV 62  
Qy 217 RGNHEDSVMARYGPRREVSKYPRNKRILAFIDEVYRVLPLGSVLSNSRLVHGGFSD 276  
Db 63 LGNHDLVLLAVGCG-----EGALKR 82

Qy 277 STS1D-LIKEIDRKGKVSIIRPLPDKTETQIDIMNSDPOATMGCYPTNLGA 335  
Db 83 SDTIEPILKRPDGKQMDMLR---AQPLIIRREGGRV---MIHAGILPQ--- 124

Qy 336 GWIGCPDVDTUNFLQRHLSVTIRSHHECKPGHFRPHEDNKLITIFSAASNTVAISNGKA- 393  
Db 125 --W-----RIAKASLSEGAEASLRGKVKVFPF---KMYCINKPAW 161

Qy 394 ---YIRANQOLMPHFWQYISANSQTKELSFKORMGIWESSALKELAWRMRDHRDELE 447  
Db 162 DEGLEYARU-----RET---VNAFTPRMALTPKNEDLTDXYSKTPKMPYLRW--- 208

Qy 448 DEFRIKYDPKOSGY-1SISHCKV---MENVTKL---GLPWRLLRKDCLKAPGTDISKUNTR 500  
Db 209 --FKAPDRQDLHTIFGHWSSLGTYTNADNVISLDTGALW---GGOLTAVN--- 254

Qy 501 TLDLDTDVLEAADG 517  
Db 255 ---LETERTIQVQAG 267

---

RESULT 27  
US-11-074-176-314  
; Sequence 34, Application US/11074176  
; Publication No. US20050250135A1  
; GENERAL INFORMATION:  
; APPLICANT: Klaenhammer, Todd R.  
; APPLICANT: Russell, William M.  
; APPLICANT: Alterman, Eric C.  
; APPLICANT: Alterman, Eric  
; APPLICANT: McAuliffe, Olivia  
; APPLICANT: Peril, Andrea Azcarate  
; APPLICANT: Peril, Andrea Azcarate  
; TITLE OF INVENTION: Nucleic Acid Sequences Encoding  
; TITLE OF INVENTION: Stress-Related Proteins and Uses Therefore  
; TITLE OF INVENTION: Stress-Related Proteins and Uses Therefore  
; FILE REFERENCE: 5051-694  
; CURRENT APPLICATION NUMBER: US/11/074,176  
; CURRENT FILING DATE: 2005-03-07  
; PRIOR APPLICATION NUMBER: 5051-694  
; CURRENT APPLICATION NUMBER: US/11/074,176  
; CURRENT FILING DATE: 2005-03-07  
; PRIOR APPLICATION NUMBER: 6/0/551,161  
; NUMBER OF SEQ ID NOS: 381  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO: 314  
; LENGTH: 746

---

RESULT 28  
US-11-074-176-54  
; Sequence 54, Application US/11074176  
; Publication No. US20050250135A1  
; GENERAL INFORMATION:  
; APPLICANT: Klaenhammer, Todd R.  
; APPLICANT: Russell, William M.  
; APPLICANT: Alterman, Eric  
; APPLICANT: Alterman, Eric  
; APPLICANT: Peril, Andrea Azcarate  
; TITLE OF INVENTION: Nucleic Acid Sequences Encoding  
; TITLE OF INVENTION: Stress-Related Proteins and Uses Therefore  
; FILE REFERENCE: 5051-694  
; CURRENT APPLICATION NUMBER: US/11/074,176  
; CURRENT FILING DATE: 2005-03-07  
; PRIOR APPLICATION NUMBER: 6/0/551,161  
; NUMBER OF SEQ ID NOS: 381  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO: 54  
; LENGTH: 749  
; TYPE: PRT  
; ORGANISM: Lactobacillus acidophilus  
; US-11-074-176-54

Query Match Similarity 2.9%; Score 98.5; DB 7; Length 749;  
 Best Local Similarity 19.6%; Pred. No. 5.9; Mismatches 217; Indels 167; Gaps 27;

Matches 114; Conservative 83; Mismatches 217; Indels 167; Gaps 27;

Qy 32 QIFONLETASBODA---ELYKPN-----DLIKHMPQAGRKNOY----- 69  
 Db 14 QVIDTCKKMMNEQDQAFVERAYEFANKAHAGOKRASGOPYIHLPTQVAGTLANGLDPD 73  
 Qy 70 --OCSAHTSYLD---DKDLYVERFGD---IWKAKLIP-----DLIKHMPQAGRKNOY----- 69  
 Db 74 TAAGFLPHDTVEDPVTNDLKECKEDYAFIIVDGUTKLNEYNSHOBFLAENHKMLIA 133  
 Qy 110 -----DVERKKGRLHPKVVALILREPAKSLKQLPNISPVSTAVSOQVTCGGDHLKUD 164  
 Db 134 MAKDRLVIMVKLADRHKMHTLQLHR-PDKQRBIASETMIDVPLABRGLI-GTIKWELE 191  
 Qy 165 DL-----LVLHKGNGLPSSSNPYVFNDFVDRGKSHLVEVLLLISLYLAPPNAV 213  
 Db 192 DMSFHFLNPEAYTRVNLMDKESQREKVI--SDTIKTKTUDLGKIVDY----- 242  
 Qy 214 FLURGNHDSVMARYGPIREVESKYPANHKLAFDEVYKMLPLGSVANRVL----- 269  
 Db 243 --GRPKH-----IYSTIKRMVKHQD---FDEYI-----DILATRVIVENR 279  
 Qy 270 -----VH-----GCFSSTSLSLJIKSIDRGKYSILRPLTDGEPPLD-KTEWQOI 313  
 Db 280 DCYAVLGAHVTHENKPMPEPKDYLAMPKNGY-QSLHITIGP---GGRPLSIQIRQM 335  
 Qy 314 FDIMWSDPOATMGCVNPNTIRGAGCWMFGDVTUDFLORHLSVYTRSHCKPGHEFMEDN 373  
 Qy :::: HEV-----AEYGAHAHWYKRGNFGNGEATSSGKLDWREILELKEDTQAGEFMKV 389  
 Db 336 KLTIFSASN-----VAIGSUKGAY---IRLANNOLMPHFWQYI 409  
 Db 390 K-SDFISDRVYVTPKGEVYELPKGSVTLDAFAIHTOVHGARGTYNKLV----- 442  
 Qy 410 SAASQTKLISKORMGIVESALKELAVMRDHRDE---DEFRKUDPKDSGYIS 464  
 Db 443 -----LDYKLRNGDVIEIMTQNTAAPSRDADMVTKTSARNKIRYFRCQDREESIR 494  
 Qy 465 HWCKWMEN-VTKUGLGPWLRLDKLAPGTDQSQVNYNRITLD 504  
 Db 495 HGEOMVANTIRDEGLVPKDMDKEHIEKULJEHFNTINTSEEL 535

---

RESULT 29 US-10-995-561-772

; Sequence 772, Application US/10995561

; Publication No. US200502720541

; GENERAL INFORMATION:

; APPLICANT: CARGILL, Michele et al.

; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH RESPONSE, METHODS OF THERAPY AND USES THEREOF

; TITLE OF INVENTION: DETECTION AND USES THEREOF

; FILE REFERENCE: C1001559

; CURRENT APPLICATION NUMBER: US/10/995, 561

; CURRENT FILING DATE: 2004-11-24

; NUMBER OF SEQ ID NOS: 85702

; SOFTWARE: FastSEQ for Windows Version 4.0

; LENGTH: 1538

; TYPE: PRT

; ORGANISM: Homo sapiens

; US-10-995-561-772

RESULT 29 US-10-995-561-772

; Sequence 772, Application US/10995561

; Publication No. US200502720541

; GENERAL INFORMATION:

; APPLICANT: CARGILL, Michele et al.

; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH RESPONSE, METHODS OF THERAPY AND USES THEREOF

; TITLE OF INVENTION: DETECTION AND USES THEREOF

; FILE REFERENCE: C1001559

; CURRENT APPLICATION NUMBER: US/10/995, 561

; CURRENT FILING DATE: 2004-11-24

; NUMBER OF SEQ ID NOS: 85702

; SOFTWARE: FastSEQ for Windows Version 4.0

; SEQ ID NO: 777

; LENGTH: 5335

; TYPE: PRT

; ORGANISM: Homo sapiens

; US-10-995-561-777

Query Match Similarity 18.9%; Pred. No. 19; Mismatches 265; Indels 292; Gaps 43; Matches 154; Conservative 103; Mismatches 265; Indels 292; Gaps 43;

Qy 9 AIFIQKWRKRQARREMORR-----CWWQIF-ONLBYASSO- 43  
 Db 524 ALLBQKW--HVSSKMEERKSKLSEALNLATEFONSLOEFINWLTABQSLNTIASPPSL 580

CURRENT FILING DATE: 2004-11-24  
NUMBER OF SEQ ID NOS: 85702  
SOFTWARE: FastSEQ for Windows Version 4.0  
SEQ ID NO: 774  
LENGTH: 5406  
TYPE: PRT  
ORGANISM: Homo sapiens

; Publication No. US20050272054A1  
; GENERAL INFORMATION:  
; APPLICANT: CARGILL, Michele et al.  
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH  
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF  
; FILE REFERENCE: C2001559  
; CURRENT APPLICATION NUMBER: US/10/995, 561  
; NUMBER OF SEQ ID NOS: 85702  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO: 779  
; LENGTH: 5415  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; US-10-995-561-779

Query Match 2.8%; Score 98; DB 6; Length 5415;  
Best Local Similarity 18.9%; Pred. No. 1.2e+02; Mismatches 265; Indels 292; Gaps 43;  
Matches 154; Conservative 103; Mismatches 265; Indels 292; Gaps 43;

QY 9 AIFIQKWRKRHQAAREMQR-----  
Db 4456 ALLRQKM--HYVSSKMERHKSKLEALNATEFONSQPIWNLIAEQLNIAASPL 4512

QY 44 -----DQAELEYKF-NDLIKHIMPO---AAGRKNQYQSAH-----VSV---- 77  
Db 4513 ILNTVLSQLIEBKHFVANBNVNAHQDQIETLDOTGNOLKFLSQKDQVLUKNLLVSQSRWB 4572

QY 78 -----LDDKODLVBERECDIVAKIE-LPIRKHNHDIDLV----- 111  
Db 4573 KVVRQSIEGRSLDARKRAKQFHAWKCLIDMWLDABSHLDLSRBLISNDPDKIKLQLSK 4632

QY 112 --TRKRGNRNHLKPVALLIREAKSLIKQ--LPNISPVSTAVSQV---TVCG-- 157  
Db 4633 HKEFQTKLGGK-QPVYDPTI--RTGRALKETKILPEDOTQKUDFLGFRDKMTVCGSKV 4689

QY 158 DLHGKDDLLVVLHKNGLPPSSSNPYVNQDFVDRGKR-----GLB 197  
Db 4690 ERHQHKLSEAU-----LFSQFMDALQALVDWLYLKVPQLAOPVHGDLD 4734

QY 198 VLLMLSLTYLAEPNAV-----FLNRGNHEDSVMMAR-----YGFIREV----- 235  
Db 4735 LVMNLMDAHKVFOKELGKRTGTQVQLKRSRE-LIENSDDTTWVKQGQLESTRWDTVC 4793

QY 236 -----ESKVPRNHRKLLAFIDEY--RWLPLGSVLSVRLVHGGFSDSISLDIJKSI 286  
Db 4794 KLSVSKQSKRNLQALKQAEVRDFTVHMLLWSEA---EQTLRPGALPDDEA-LSLII 4848

QY 287 DRGKVSVIRPPLTDGEPLDKTEQFDIMSDPQATMG-----CVPNTLRGAGWFG 340  
Db 4849 D-----THKEFKMKVEKRV-DV--NSAVANGEVILAVCHPCDCITKHWI- 4891

QY 341 PDVTUNFLQRHLSTVIRHECKINGHFPMDNKITITFASNYTAGSNSKGAYIRLNO 400  
Db 4892 -----TIRARFEELVTLWAK-----OHOORLETALSE-----LVANAE 4924

QY 401 IMPHPVQYISAASQT-----KRLSFKORMGIVESALKELAVRMRDHD 444  
Db 4925 LLBELLAWIOWAQWATTTLQDQEPIPONIDRVAL-----LAHOTNEEMTRKOPYD 4977

QY 480 -WLRLARDKLQAPGTOSOKVNNTRDUDTIVLRAAD-----GMSYMDALVANKASLV 532  
Db 5038 VMLLALER-----ORQKLN-----DALDRLEELKEFANPDFDWVKKMRWWMHKKSRW 5086

QY 533 AIFNIDADSGETLDEFTAID-LVVAHMPGYSKAELMEKCRMMDLNGDSKUDLNEP 591  
Db 5087 DPFRRIDKQDGKTRQBP---IDGIAKRP--TTKLMNTAVADIFORGDGSYTDPF 5141

RESULT 33  
US-10-995-561-775  
Sequence 775, Application US/10995561  
Publication No. US20050272054A1  
GENERAL INFORMATION:  
APPLICANT: CARGILL, Michele et al.  
TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF  
FILE REFERENCE: C2001559  
CURRENT APPLICATION NUMBER: US/10/995, 561  
NUMBER OF SEQ ID NOS: 85702  
SOFTWARE: FastSEQ for Windows Version 4.0  
SEQ ID NO: 775  
LENGTH: 5464  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-995-561-775

Query Match 2.8%; Score 98; DB 6; Length 5464;  
Best Local Similarity 18.9%; Pred. No. 1.2e+02; Mismatches 265; Indels 292; Gaps 43;  
Matches 154; Conservative 103; Mismatches 265; Indels 292; Gaps 43;

QY 9 AIFIQKWRKRHQAAREMQR-----  
Db 4456 ALLRQKM--HYVSSKMERHKSKLEALNATEFONSQPIWNLIAEQLNIAASPL 4512

QY 44 -----DQAELEYKF-NDLIKHIMPO---AAGRKNQYQSAH-----VSV---- 77  
Db 4513 ILNTVLSQLIEBKHFVANBNVNAHQDQIETLDOTGNOLKFLSQKDQVLUKNLLVSQSRWB 4572

QY 78 -----LDDKODLVBERECDIVAKIE-LPIRKHNHDIDLV----- 111  
Db 4573 KVVRQSIEGRSLDARKRAKQFHAWKCLIDMWLDABSHLDLSRBLISNDPDKIKLQLSK 4632

QY 112 --TRKRGNRNHLKPVALLIREAKSLIKQ--LPNISPVSTAVSQV---TVCG-- 157  
Db 4633 HKEFQTKLGGK-QPVYDPTI--RTGRALKETKILPEDOTQKUDFLGFRDKMTVCGSKV 4689

QY 158 DLHGKDDLLVVLHKNGLPPSSSNPYVNQDFVDRGKR-----GLB 197  
Db 4690 ERHQHKLSEAU-----LFSQFMDALQALVDWLYLKVPQLAOPVHGDLD 4734

QY 198 VLLMLSLTYLAEPNAV-----FLNRGNHEDSVMMAR-----YGFIREV----- 235  
Db 4735 LVMNLMDAHKVFOKELGKRTGTQVQLKRSRE-LIENSDDTTWVKQGQLESTRWDTVC 4793

QY 236 -----ESKVPRNHRKLLAFIDEY--RWLPLGSVLSVRLVHGGFSDSISLDIJKSI 286  
Db 4794 KLSVSKQSKRNLQALKQAEVRDFTVHMLLWSEA---EQTLRPGALPDDEA-LSLII 4848

QY 287 DRGKVSVIRPPLTDGEPLDKTEQFDIMSDPQATMG-----CVPNTLRGAGWFG 340  
Db 4849 D-----THKEFKMKVEKRV-DV--NSAVANGEVILAVCHPCDCITKHWI- 4891

QY 341 PDVTUNFLQRHLSTVIRHECKINGHFPMDNKITITFASNYTAGSNSKGAYIRLNO 400  
Db 4892 -----TIRARFEELVTLWAK-----OHOORLETALSE-----LVANAE 4924

QY 401 IMPHPVQYISAASQT-----KRLSFKORMGIVESALKELAVRMRDHD 444  
Db 4925 LLBELLAWIOWAQWATTTLQDQEPIPONIDRVAL-----LAHOTNEEMTRKOPYD 4977

QY 341 PDVTUNFLQRHLSTVIRHECKINGHFPMDNKITITFASNYTAGSNSKGAYIRLNO 400  
Db 4892 -----TIRARFEELVTLWAK-----OHOORLETALSE-----LVANAE 4924

QY 401 IMPHPVQYISAASQT-----KRLSFKORMGIVESALKELAVRMRDHD 444  
Db 4925 LLBELLAWIOWAQWATTTLQDQEPIPONIDRVAL-----LAHOTNEEMTRKOPYD 4977

QY 445 ELEDEFRK-YDPKDSGYI-SISHWCKMENVTKLGLP----- 479  
Db 4978 RVTKTYKRNIEPTHAPEKFSRSGGRKSISQTPPPMPLSOSBEAKPRINQLSARWQ 5037

QY 480 -WLRLARDKLQAPGTOSOKVNNTRDUDTIVLRAAD-----GMSYMDALVANKASLV 532  
Db 5038 VMLLALER-----ORQKLN-----DALDRLEELKEFANPDFDWVKKMRWWMHKKSRW 5086

QY 533 AIFNIDADSGETLDEFTAID-LVVAHMPGYSKAELMEKCRMMDLNGDSKUDLNEP 591  
Db 5087 DPFRRIDKQDGKTRQBP---IDGIAKRP--TTKLMNTAVADIFORGDGSYTDPF 5141

QY 4978 RVTKTYKRNIEPTHAPEKFSRSGGRKSISQTPPPMPLSOSBEAKPRINQLSARWQ 5037

QY 480 -WRLRDKLAPGTDQSOKVNRTLUDLTDVILEAAD----GMSYDALYANKASLV 532  
Db 5038 WMLALER----ORKIN----DAJDLRBLKEPANFDWVRKKTMRWMNHKSVM 5086  
QY 533 AIFNITDADNSGBTIDTFETAD-LLYAHMPGAYSKARMLEKCRMMDLNGDKVLDNF 591  
Db 5087 DFFRKDKDQDGKLTROFR--IDGILASKFP--TTKLEMATAVADIFDRDG3YIDYBF 5141  
QY 592 LEAPRLS-DLHRHQODENIRRSTGRPEVAKTA 624  
Db 5142 VAALHPNKDVAIRPTTADKIEDEVTRQVACKCA 5175

RESULT 34  
US-10-995-561-776  
; Sequence 776, Application US/10995561  
; Publication No. US20050272054A1  
; GENERAL INFORMATION:  
; APPLICANT: CARGILL, Michele et al.  
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH  
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF  
; FILE REFERENCE: CL0011559  
; CURRENT APPLICATION NUMBER: US/10/995,561  
; CURRENT FILING DATE: 2004-11-24  
; NUMBER OF SEQ ID NOS: 85702  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 776  
; LENGTH: 5935  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; US-10-995-561-776

Query Match Score 2.8%; Score 98; DB 6; Length 5935;  
Best Local Similarity 18.9%; Pred. No. 1.4e+02; Mismatches 154; Conservative 103; Mismatches 265; Indels 292; Gaps 43; Length: 5935

QY 9 AIFIQKWKHQRARREMOR-----CWNQIF-ONLYEVASEQ-- 43  
Db 4958 ALLBQKWW--HVSSKMEERKSKEBALNIAFQNSLOFPIWNLTLABQSLNTAAPPSSL 5014

QY 44 -----DOALEYKEF-NDLKHMQ---AAGRKXQGSAH-----VSU---- 77  
Db 5015 ILNTVLSQEHEKYTFANEYNAHRDQIBLDTGNOLKFLSKQDOWVLKNNLUVSQSRWE 5074

QY 78 -----LDDKDLVEEFDIVAKIE-LPIRKHHIDLLDV----- 111  
Db 5075 KWNVBSIERRSDDARKKAQFHAWKLDIDWEDASHBLSBISNDPKXQLISK 5134

QY 112 ---FRKKRKRNLHFKYVALLTREAKSLQ--LPNISPVSTAVSQY----TVCG--- 157  
Db 5135 HKFBQFTLGK-QPVDTT--RTGRALKETKTLUPDFTQKDNPLGEVRDKWDIVCGKV 5191

QY 158 DLHKGDDLWVWHNGUBSSSNIVYFVKDPVUGKR-----GLB 197  
Db 5192 BROHLERAL-----LFGQFMALQALDYLWVLYKVPOLAEQPVHGQD 5236

QY 198 VLLLILSYLAEPNAV-----PLNRGNHEDSVNAR-----YGFIREV----- 235  
Db 5237 LWMLMDAHLWVFOLEGKETGTIVOKLRSGG--LIENSBDTWTWKGQJQELSSTRWTC 5295

QY 236 -----ESKPKNRKRILAFIDEY--RWLPLGSVLSRVLVHGGSDSTSLLDKSI 286  
Db 5296 KLSVSKQSLRQEQLQAEVFRDVTWMLLWSEA---EQTLRFRGALEPDTEA-LOSLI 5350

QY 287 DRGKVSLIRPLPTEPDKTKEQLEIMMSPQATW-----CYPMTLRGAGWFG 340  
Db 5351 D-----THKEFMKVKEKRV-DV---NSAVANGEVILAVCHDCTTICKWI- 5393  
QY 341 PDVTDNFLQHRLSYVIRSHCECKNGHEFWHDNKITIPFSASNYAIGSNKGAYIRLNO 400  
Db 5394 -----TIRARESEVLTWAK-----CQHORLLETALB-----LVAAN 5426

QY 401 LMPHFVQVTSASQQT-----KRLSPKFORMGTVESSALKELAVRMRIIRD 444  
Db 5427 LLERBLANIQWAEATTIORDQEPIPONIDRVKAL-----IAEHQPFMEBTRKQPDVD 5479

QY 445 BLDEPERK--YDKDGYI-SISHWCIVMENUTKGUL----- 479  
Db 5480 RVTKTYKRNKTBPHAPTEKSRSGGKRLSISOPTPPMPILSOEAKNPRINOLSARWQO 5539

QY 480 -WRLRDKLAPGTDQSOKVNRTLUDLTDVILEAAD----GMSYDALYANKASLV 532
Db 5644 VAALHPNKDVAIRPTTADKIEDEVTRQVACKCA 5677

RESULT 35  
US-11-156-084-115  
; Sequence 115, Application US/11156084  
; Publication No. US2006010515A1  
; GENERAL INFORMATION:  
; APPLICANT: Monsanto Technology LLC  
; TITLE OF INVENTION: Controlled expression of cytokinin biosynthetic genes leads to FILE REFERENCE: (38-21)  
; CURRENT APPLICATION NUMBER: US/11/156, 084  
; CURRENT FILING DATE: 2005-06-17  
; NUMBER OF SEQ ID NOS: 364  
; SOFTWARE: Patentin version 3.2  
; SEQ ID NO 115  
; LENGTH: 467  
; TYPE: PRT  
; ORGANISM: Medicago truncatula  
; US-11-156-084-115

Query Match Score 2.8%; Score 97.5; DB 7; Length 467;  
Best Local Similarity 18.5%; Pred. No. 3; Mismatches 81; Conservative 54; Mismatches 140; Indels 163; Gaps 18; Length: 467

QY 36 NLBYASQDQAELYKPFDLI-KHMMAQAGRKNQY QGSAAHVSVLDDK-DDLVEE---- 88  
Db 83 NLBFTAKTUFTRDSLAPIDEILSRBNHLPVIGGMYVYQALVSQLPSQFLSDSTDLSRSTVLD 142

QY 89 -----GDTUNAKTELPLTRKNHIDLLIVFVKKGRGNRHP-----KVALITR----- 130  
Db 143 SPGTGFDTNFVAENDSSNSNDLKOI-DPVANRTHPNNRKINGQVISLISRTGVLP 201

QY 131 -----BAKSKLKPLNTPVSTAVSQY---VCGDHLGKDLLVWLRKONGLUPSSNPY 182  
Db 202 KVFGQQAERKKGQDADNRYACCPISVDSLPLTLDVRYEVQDMMADYTRGLQAT--- 258

QY 183 VFNGDFDVRGKGKLEVLILLSLYLAPPNAFLVNRGHNEDSVNARYGFIREVESKF-- 240  
Db 259 -----GVERPFDL-----RTSFVNKINGREGELJDG----SSLNSPLP 295

QY 241 -----RNUHKRILAPIDEYR----- 256  
Db 296 DGNLMAMWLSFSPTKSTLLEDALEKQVNTLVRQKRMLSRLETLFGWDIHVYSTE 355

QY 257 -----LPLSVTNRVLYHGGFDSTSLLDKSISLDRGKVSVIRPLPDTGPDKT 308  
Db 356 SILSKSDIUNRNOVSEATKIVTIVBNSLSTFCMSNGTKLQD 406

QY 379 EWOQIPDIMSPLQATMCVPUVLRGAGWFGDPTVNLQH----- 351  
Db 407 -WTO-----YTCHKACNDRTVGHW-----BOHIGRGRKRKISSLSKA 446

352 -RLSYVIRSHECKPNH 368

```

Db      447 KGKSFV--EKKCPSHEE 462
Db      305 EHYDNPTIES-----GLDNGSLR-LFTYM-----DIRKY---DAGVIBAG 342
Qy      465 HKCKMENVTKGGLPMLRLDKLAPGTDQSQVNNTLTLADTDVLEBADGMSWMDAL 524
Db      343 LWSLFRHTI-----PPGMPEFOSEGHCCTLCBL-----BAL 373
Qy      525 YANKASIVAFNIDADDNSGETILDEFFETADLUVAHMPGYSKARMLEKCRMMDING-D 583
Db      374 RAHKPSGIVHFAV-----LLHAHLAGRGIRLRHFRKGKEMKLAYD 414
Qy      584 GKVDLNFLEATRLSDIARKBQQ 606
Db      415 DDPDFN--FQEQQ---YLKEEQ 431
Db      37 RESULT 37
; Sequence 2, Application US/10118590
; Publication No. US20050277761A1
; GENERAL INFORMATION:
; APPLICANT: KENNETH RHODES, MARIA BETTY, HUAI-PING LING, AND FRANK AN
; TITLE OF INVENTION: POTASSIUM CHANNEL INTERACTORS AND USES THEREFOR
; FILE REFERENCE: MINI-070
; CURRENT APPLICATION NUMBER: US/10/118 590
; CURRENT FILING DATE: 2002-04-08
; PRIOR APPLICATION NUMBER: US/09/298, 731
; PRIOR FILING DATE: 1999-04-23
; NUMBER OF SEQ ID NOS: 47
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO: 2
; LENGTH: 216
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-118-590-2
Query Match          2.81! Score 97; DB 6; Length 216;
Best Local Similarity 20.8%; Pred. No. 1.2; Mismatches 69; Indels 88; Gaps 9;
Matches 50; Conservative 33; Mismatches
Qy      424 MGIVBSALKELAVRMPDHDLERDKYKPDQDSGYISISHWMKOMENVTKGGLPWRLL 483
Db      5 MGIVFSSTQTKO---RPSKDKEDR-----LEMTHVCHPREGLQLQBLACTNPFT 49
PRIOR APPLICATION NUMBER: 60/059122
PRIOR FILING DATE: 1997-09-17
PRIOR APPLICATION NUMBER: 60/059184
PRIOR FILING DATE: 1997-09-17
PRIOR APPLICATION NUMBER: 60/059263
PRIOR FILING DATE: 1997-09-18
PRIOR APPLICATION NUMBER: 60/059352
PRIOR FILING DATE: 1997-09-19
PRIOR APPLICATION NUMBER: 60/059588
PRIOR FILING DATE: 1997-09-19
; Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 550
SEQ ID NO: 190
LENGTH: 613
TYPE: PRT
ORGANISM: Homo Sapien
US-10-111-826A-190
Query Match          2.81! Score 97; DB 6; Length 613;
Best Local Similarity 21.4%; Pred. No. 5.3; Mismatches 103; Indels 113; Gaps 17;
Matches 69; Conservative 38; Mismatches
Qy      306 DKTWOQIQT-MWSDPQATMGCVPLLAGAGVWPGDVTNQFRHLSVYRSHKCP 364
Db      200 DTTWQCMRKPVVKPHVHVVQHG-----HESVHHLINY----QCSN 244
N-----CHEPMEDN--KIIITPSASNYAIGSNKAY----IRNQJMPH---- 405
Db      245 NFNDVSYLESGHECYPHPNMPPAFLCTEVTFPAWAGGGPSYPPHVGSLGTPDPHVLL 304
Qy      406 -VQITSAASOTKRLSKQRMGIVESSALKEVALVRMRDRHDBELDFRKYDPKSGYTSIS 464
; Sequence 6, Application US/10118590
; Publication No. US20050277761A1
; GENERAL INFORMATION:
; APPLICANT: KENNETH RHODES, MARIA BETTY, HUAI-PING LING, AND FRANK AN
; TITLE OF INVENTION: POTASSIUM CHANNEL INTERACTORS AND USES THEREFOR
; FILE REFERENCE: MINI-070
; CURRENT APPLICATION NUMBER: US/10/118 590
; CURRENT FILING DATE: 2002-04-08
; PRIOR APPLICATION NUMBER: US/09/298, 731
; PRIOR FILING DATE: 1999-04-23
; NUMBER OF SEQ ID NOS: 47
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO: 6
; LENGTH: 216
; TYPE: PRT
RESULT 38
US-10-118-590-6
; Sequence 6, Application US/10118590
; Publication No. US20050277761A1
; GENERAL INFORMATION:
; APPLICANT: KENNETH RHODES, MARIA BETTY, HUAI-PING LING, AND FRANK AN
; TITLE OF INVENTION: POTASSIUM CHANNEL INTERACTORS AND USES THEREFOR
; FILE REFERENCE: MINI-070
; CURRENT APPLICATION NUMBER: US/10/118 590
; CURRENT FILING DATE: 2002-04-08
; PRIOR APPLICATION NUMBER: US/09/298, 731
; PRIOR FILING DATE: 1999-04-23
; NUMBER OF SEQ ID NOS: 47
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO: 6
; LENGTH: 216
; TYPE: PRT

```

; ORGANISM: Mus musculus  
; US-10-118-590-6

Query Match Similarity 2.8%; Score 97; DB 6; length 216;  
Best Local Similarity 20.8%; Pred. No. 1.2.;  
Matches 50; Conservative 33; Mismatches 69; Indels 88; Gaps 9;

Qy 424 MGIVESSALKELAVERMRDHDELEFRKDPKGYSVISHWCKMVENTKGLPWRUJ 483  
Db 5 MGTFFSLLQTQK---RPSKDKEB-----LEMMVCHRPPEGILQOLEAQTNFT 49

Qy 484 RDKARGTSDOKUNRTIDLDLTVILBABAEGMSVMALYA---NKSILVA-IRN 536  
Db 50 KREL----QVLTGFKNECPGVNB----ETEKQIYAOFPFGHASTAYHEN 96

Qy 537 IIDADNSGETLDERETATDOLV-----AMPGAVSKAEMLEKCR--- 576  
Db 97 AFDITOTGSVKREDFVTALISLURGVHKLWTFNLVIDNKDGYINKESMDVIAID 156

Qy 577 -----MHDLDNGDKVULNEFRAFRSLDRKKEQDENTIR 612  
Db 157 MMGKYTYVPVLIKEDTPRHVDVFQKMDKNKGIVLDEFESC-----QEDDNIMR 207

RESULT 39  
US-10-878-556A-161  
; Sequence 161, Application US/10/878556A  
; Publication No. US20050266399A1  
; GENERAL INFORMATION:  
; APPLICANT: Hoffmann La Roche Inc.  
; TITLE OF INVENTION: Hcv regulated protein expression  
; FILE REFERENCE: 21762  
; CURRENT APPLICATION NUMBER: US/10/878,556A  
; NUMBER OF SEQ ID NOS: 199  
; SEQ ID NO 161  
; LENGTH: 814  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; PUBLICATION INFORMATION:  
; DATABASE ACCESSION NUMBER: SW\_hum/1f39\_human  
; DATABASE ENTRY DATE: 1997-11-01  
; US-10-878-556A-161

Query Match Similarity 2.8%; Score 97; DB 6; length 814;

Best Local Similarity 20.8%; Pred. No. 1.2.;  
Matches 73; Conservative 55; Mismatches 119; Indels 124; Gaps 16;

Qy 185 NGDF---YRGKRGKHLVLILLSLVLAAPMAVPLNRGNGHEDSVNARYGPIRVESKP 240  
Db 522 NGDYLCVKVDRTPKGQGVVT-----NPFIFRKREKOPVPUVEMK-- 562

Qy 241 RHKRLAPLDEVKVLPLASVLSVLRIG-----GFDSTSDLISIDRK 290  
Db 563 ---ETTIAF----AWEPNG---SKPAVINGEARPRISVSYHVNNGKELIKNFDK-- 608

Qy 291 YVSILRPLTTEGPLDKTEWQIFDMWSIDPOATMCCVPUWTLRAGVAGWFCDDVUNFLQR 350  
Db 609 -----QANTIFMS PQSQFVIALRSRSMQALAFDTSCTVNN 647

Qy 351 HRLSYVIRSHCKPKNGHEFMNDKNTIFS----ASNYAIGSNKGAYRLNNOLMPHPVQYTAASQTKRIS 405  
Db 648 IAETHMASDVEWDPTG----RYVTSVSWSHKDNAIWLWTFQGRLLQKNNK--DRP 699

Qy 406 VOY-----ISAQSOKRIS-----FKQMGIVSSALKELAVERMRDHDELE 448  
Db 700 CQLLWRPRPRTPLSQQIKQTKDLCKKYSKFEQDRLSOKSKASKLVERRR--TMME 755

Qy 449 EFRKYPDKSGYISISHWCKMVENTKGLGEWPLWLRDKLAPGTDQKVNKN-----RTV 502  
Db 756 DFRKY-----RKMADELYMQKNERLLELGVOPTBLDSNVWWEBET 799

Qy 503 DLLDTDVIL 511  
Db 800 EFFVTBII 808

RESULT 40  
US-11-115-639-50

Sequence 50, Application US/11115639  
; Publication No. US20050282242A1

; GENERAL INFORMATION:  
; APPLICANT: Rothstein, David  
; APPLICANT: Murphy, Christopher  
; APPLICANT: MacNeil, Ian  
; TITLE OF INVENTION: SCREENING ASSAYS FOR ANTIMICROBIAL  
; TITLE OF INVENTION: AGENTS  
; FILE REFERENCE: 50150/075003  
; CURRENT APPLICATION NUMBER: US/11/115,639  
; CURRENT FILING DATE: 2005-04-27  
; PRIORITY NUMBER: 60/566,858  
; PRIORITY NUMBER: 60/565,679  
; PRIORITY NUMBER: 2004-04-27  
; NUMBER OF SEQ ID NOS: 55  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 50  
; LENGTH: 1184  
; TYPE: PRT  
; ORGANISM: L. monocytogenes  
; US-11-115-639-50

Query Match Similarity 2.8%; Score 97; DB 7; Length 1184;

Best Local Similarity 20.8%; Pred. No. 1.5;  
Matches 157; Conservative 88; Mismatches 257; Indels 252; Gaps 41;

Qy 16 YRHHQARREMRQRMQRNWQIQFQNLEYAS---EQDQBLKPKFNDLKHMPPQAGRNQYQDQA 73  
Db 11 YGRHTRRSFAR-----ISRVBLPNLIBQIOTASYQWFLDEGLERMFIDSPIBDFAGNL 65

Qy 74 HVSLVDDKODLVEERGDIWAKBLPIRKHIDLVIDPKKRRGNRHPKVVALIREAA 133  
Db 66 SLEPFD-----YDGERPKYVSVEREN-----RQANYAALPRLVKLRL--- 101

Qy 134 KSLKLPNTPSPVSTAVSQQVTCED---LNGKDDDLVUVLHNG-----LPSSSNPV 183  
Db 102 -----INKEΤGEVKDQEVMGDPPLMTENGTFPI---NGAERVIVSOLVRSPGVY- 148

Qy 184 FNQGPFDKRGKHLVLILLSLVLAAPMAVPLNRGNGHEDSVNARYGPIRVESKP 228  
Db 149 FNGKLDKONGKKG-----PGSTVIPNRGAWLBYETDKDVYHVRIDTRKLPUV 195

Qy 229 -----YGPIREV---ESKIPRN-----HKLAPIDEVYRMLPLG---SV 262  
Db 196 TULLRALAGGSDQDILDLGNDNTLYRNLTEKDWTDNASKALL---EYERLNPGRPPTV 251

Qy 263 LNSRVTIVGGFSSTDSDLIKSTDGRK-----YVSLIRPLTD----- 301  
Db 252 DNARSLSLV-SRFPDPKRIDL-ASVGRKINKKGLKLNRLPNQTLATELVDPETGBIAISK 309

Qy 302 GEPDJKTEWQIQIFDMSDPOATMCCVPUWTLRAGVAGWFCDDVUNFLQR 360  
Db 310 GDLDRRNLDQI-----IPLENLGVGFTLRP--TDGVMEDSVLVSQSKY 353

Qy 361 ECKNGHEPMNDKNTIFSASNYAIGSNKGAYRLNNOLMPHPVQYTAASQTKRIS 419  
Db 354 --APNDEB-----KEINI-----IGN---AVIEN-----VKHTPSDIIISISY 388

Qy 420 -FKQMGIVSSALKELAVERMRDHDELEFRKDPKGYSVISHWCKV-----MB 471  
Db 389 PFNLJHGVDIDPDLHGRRLSUGVQNR-----IGLSMERMVRBRMSIQ 439

Qy 472 NVTKLGPWRL-----LDRKLAPGTDQKVNKN-----E 514  
Db 440 DMFTI-TPOOLINRPPVVASIKERPGSSOLSOFNQTNPLDELTYKRRSLALGPGGLTR 498

OY 515 ADGMSVMDALYANAKSL---.VAIFNTIDADNSGETLDBF-----E 552  
Db 499 RAGYEVRDWYSHYGRMCPIETPFGPNIGLENTI---.SSPAKUNKFOPIETPYRYDPE 554  
OY 553 T-----AIDLIVAINMPGAVSKASNLEKCRMMDINGDKVLDNEFLAFLRSOLHRSQD 607  
Db 555 TNRVTIKDLYLTADEBDNVVQAQNSK---.LDBQGTFTEBEMARPRSENJ---.AVEK 606  
OY 608 ENIRRSTGRPSVAKTADPPVTLADKISKNTLV 641  
Db 607 ERIDYMDYSPKQVSVATACIPFLENDISNRAIM 640

Search completed: January 20, 2006, 20:04:22  
Job time : 37 secs

GenCore version 5.1.6  
 Copyright (c) 1993 - 2006 Compugen Ltd.

## OM protein - protein search, using SW model

Run on: January 20, 2006, 19:39:22 ; Search time 40 seconds  
 (without alignments)  
 1589.982 Million cell updates/sec

Title: US-09-463-733-1  
 Perfect score: 3442  
 Sequence: 1 MDENATRAIFIQKWRHQQ... VEHIDIDPDCESKVDPKKS 661

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0  
 Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
 Maximum Match 100%  
 Listing first 45 summaries

Database : PIR-80;\*

1: pir1:\*

2: pir2:\*

3: pir3:\*

4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3442	100.0	661	1	A42287 phosphoprotein pho
2	1317	38.3	707	2	T42239 probable phosphor hypothetical prote hypothetical prote
3	1305	37.9	722	2	T4072
4	593.5	17.2	704	2	T41614 phosphoprotein pho
5	569	16.5	499	1	A55346 phosphoprotein pho
6	567.5	16.5	498	1	S2570 phosphoprotein pho
7	558	16.2	526	1	T45058 phosphoprotein pho
8	557.5	16.2	533	2	B84858 phosphoprotein pho
9	541	15.7	479	1	T4576 phosphoprotein pho
10	519	15.1	473	1	T40391 phosphoprotein pho
11	479.5	13.9	513	1	S52571 phosphoprotein pho
12	444	12.9	413	2	T51611 phosphoprotein pho
13	441	12.8	326	2	T09959 phosphoprotein pho
14	436.5	12.7	314	2	S52371 phosphoprotein pho
15	435.5	12.7	327	2	A25500 phosphoprotein pho
16	434	12.6	729	2	T0913 phosphoprotein pho
17	433.5	12.6	326	2	T05544 phosphoprotein pho
18	429.5	12.5	323	2	A22549 phosphoprotein pho
19	429	12.5	322	2	B22550 phosphoprotein pho
20	428	12.4	318	2	S20882 phosphoprotein pho
21	426.5	12.4	317	2	T05594 phosphoprotein pho
22	425	12.3	401	1	T21288 phosphoprotein pho
23	424.5	12.3	312	2	S31086 phosphoprotein pho
24	421	12.2	323	2	T05550 phosphoprotein pho
25	420.5	12.2	317	2	T03304 phosphoprotein pho
26	420	12.2	291	1	T21322 phosphoprotein pho
27	419.5	12.2	321	2	S31088 phosphoprotein pho
28	418.5	12.2	310	1	T25993 phosphoprotein pho

## ALIGNMENTS

30	418.5	12.2	312	2	S32595 phosphoprotein pho
31	418.5	12.2	312	2	S34264 phosphoprotein pho
32	418.5	12.2	325	2	T09547 phosphoprotein pho
33	413.5	12.0	401	2	T18936 phosphoprotein pho
34	413	12.0	302	1	R421A phosphoprotein pho
35	412	12.0	316	2	S26225 phosphoprotein pho
36	411.5	12.0	308	1	PABY3 phosphoprotein pho
37	411.5	12.0	333	1	T16476 phosphoprotein pho
38	409.5	11.9	382	1	T19701 phosphoprotein pho
39	409	11.9	316	2	S29317 phosphoprotein pho
40	408	11.9	312	2	S31089 phosphoprotein pho
41	408	11.9	322	1	S31087 phosphoprotein pho
42	407.5	11.8	323	2	T176572 phosphoprotein pho
43	407.5	11.8	337	1	C32550 phosphoprotein pho
44	406.5	11.8	337	1	T176573 phosphoprotein pho
45	405.5	11.8	323	1	S35699 phosphoprotein pho

QY 1 MDENATRAIFIQKWRHQQ... VEHIDIDPDCESKVDPKKS 661  
 QY 121 HKVYVALLREAKSLQLRNPSPVSTAVSQVTGDKLQFLWVHKNGLPSSN 180  
 QY 121 HPKYVALLREAKSKQLRNPSPVSTAVSQVTGDKLQFLWVHKNGLPSSN 180

Db	181	PYVFGDFDUFGRKRSRVEVLLISLKLAFPRAVFLRGHEQHEDSVIMARYFIREBSPK	240	Qy
Qy	241	RNHKRILAFIDEVVWPLPLASVNLNRYLIVGGFDSTSLIKSDIDRKVYSIRPLT	300	Db
Db	241	RNHKRILAFIDEVVWPLPLASVNLNRYLIVGGFDSTSLIKSDIDRKVYSIRPLT	300	Qy
Qy	301	DKEPKDITKESQIPIFJMASPQATMCVPTILRGAGWKGFDVTNFLQRRLSIVRSH	360	Db
Db	301	DKEPKDITKESQIPIFJMASPQATMCVPTILRGAGWKGFDVTNFLQRRLSIVRSH	360	Qy
Qy	361	ECKPQHGHEPMKDNTKITTFSASNYYAIGSNKGAYTRANNOLMPHFVYISAASQTRLSF	420	Db
Db	361	ECKPQHGHEPMKDNTKITTFSASNYYAIGSNKGAYTRANNOLMPHFVYISAASQTRLSF	420	Qy
Qy	421	KORMGIVESALKELAVMRDHDRDLEDERKYPDKDSCYISINCKMENVTUGLGLP	480	Db
Db	421	KORMGIVESALKELAVMRDHDRDLEDERKYPDKDSCYISINCKMENVTUGLGLP	480	Qy
Qy	481	RLLRKQLAPETDSOKUNRTLDLTDVILAKSKNTLVHEDPDTQESKVDPKK	540	Db
Db	481	RLLRKQLAPETDSOKUNRTLDLTDVILAKSKNTLVHEDPDTQESKVDPKK	540	Qy
Qy	541	DNSGTTLDPEATAIDLVAHMGPAGSKAEMLKCRMMDJNGDKYDLNEFLLEARFLSDL	600	Db
Db	541	DNSGTTLDPEATAIDLVAHMGPAGSKAEMLKCRMMDJNGDKYDLNEFLLEARFLSDL	600	Qy
Qy	601	HRKEDODENIRRSTGRPSVAKTADPVTIADKSKNTLVHEDPDTQESKVDPKK	660	Db
Db	601	HRKEDODENIRRSTGRPSVAKTADPVTIADKSKNTLVHEDPDTQESKVDPKK	660	Qy
Qy	661	S 661		Db
Db	661	S 661		Qy
<b>RESULT 2</b>				
T42239	probable phosphoprotein phosphatase (EC 3.1.3.16) - <i>Caenorhabditis elegans</i>			
C;Species:	<i>Caenorhabditis elegans</i>			
C;Date:	03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 12-Jul-2004			
C;Accession:	T42239			
R;Sherman, P.M.; Sun, H.; Macke, J.P.; Williams, J.; Smallwood, P.M.; Nathans, J.				
Proc. Natl. Acad. Sci. U.S.A. 94, 11639-11644, 1997				
A;Title:	Identification and characterization of a conserved family of protein serine/threonine kinase repeats in <i>Caenorhabditis elegans</i>			
A;Reference number:	Z22115; MID:97471020; PMID:932663			
A;Accession:	T42239			
A;Status: preliminary; translated from GB/EMBL/DDJB				
A;Molecule type: mRNA				
A;Residues: 1-707 <SHE>				
A;Cross-references: UNIPROT:001921; UNIPARC:UPI0000060587; EMBL:AF023454; MID:932586409; PIDN:AACT113				
C;Keywords: EF hand; phosphotriester hydrolase; serine/threonine-specific phosphatase				
Query Match 38.3%; Score 1317; DB 2; Length 707;				
Best Local Similarity 42.1%; Pred. No. 5.3e-75; Matches 266; Conservative 116; Mismatches 200; Indels 50; Gaps 12;				
Query Match 37.9%; Score 1305; DB 2; Length 722;				
Best Local Similarity 41.6%; Pred. No. 3.1e-74; Matches 266; Conservative 117; Mismatches 201; Indels 56; Gaps 13;				
Query Match 37.9%; Score 1305; DB 2; Length 722;				
Best Local Similarity 41.6%; Pred. No. 3.1e-74; Matches 266; Conservative 117; Mismatches 201; Indels 56; Gaps 13;				
Qy	6	IRAAFIQKWRHARREMRCCRQHQIQLQYASBODQABLYKEFNDLKHMPPQAGR	65	Qy
Db	83	IKSALLIQKWRRCLEARRARRAWQIFIALBEYQHGQDQKLYDFADVIRAMALENGK	142	Db
Qy	66	KNQYQG--SAHVSVLDD--KDLVVERFGDTNAKIE-----LPIRK 102	149	Qy
Db	143	GGVENCNRNPLMSALSHYAKPSLMDSEGBTWKMMLDTSPNTWDIDRNKRPPLSLPLDK	202	Db
Qy	103	NHDIDLIDVFRKKRGRMLHPKVALTIREAKSLKOLPNISPVSTAVSQQTVCGDHLGK	162	Qy
Db	203	PQVAKMIEAF--KVNKLHPKVLMLHEARKKIPKAMPSPVSRISTISNQTCIGDHLGK	260	Db
Qy	163	LDDLUVTLHKGLPSSNPYVNGDFDRGKGLYKULLLISLYLAFFPNVFLRNENHED	222	Qy
Db	261	FDDLCILYKNGYPSVDNPYFNGDFDRGKGLYKULLLISLYLAFFPNVFLRNENHED	320	Db
<b>RESULT 3</b>				
T34072	hypothetical protein F23H11.8 - <i>Caenorhabditis elegans</i>			
C;Species:	<i>Caenorhabditis elegans</i>			
C;Date:	29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 12-Jul-2004			
C;Accession:	T34072			
R;Rohlfing, T.; Wohlmann, P.				
Submitted to the EMBL Data Library, May 1997				
A;Description: The sequence of C. elegans cosmid F23H11.				
A;Reference number:	Z21470			
A;Accession:	T34072			
A;Status: preliminary; translated from GB/EMBL/DDJB				
A;Molecule type: DNA				
A;Residues: 1-722 <ROK>				
A;Cross-references: UNIPROT:001921; UNIPARC:UPI0000175983; EMBL:AF003389; PIDN:AACT113				
A;Experimental source: strain Bristol N2; clone F23H11				
C;Genetics:				
A;Gene: CRSP:F23H11.8				
A;Map position: 3				
A;Introns: 91/1; 139/3; 333/3; 579/1; 674/3				
C;Superfamily: serine/threonine protein phosphatase with EF-hands; calmodulin repeat homologous domain				
C;Keywords: EF hand				
Query Match 37.9%; Score 1305; DB 2; Length 722;				
Best Local Similarity 41.6%; Pred. No. 3.1e-74; Matches 266; Conservative 117; Mismatches 201; Indels 56; Gaps 13;				
Qy	4	NATRAIFIQKWRHARREMRCCRQHQIQLQYASBODQABLYKEFNDLKHMPPQAGR	57	Qy
Db	90	STIKSALLIQKWRRCLEARRARRAWQIFIALBEYQHGQDQKLYDFADVIRAMALENGK	149	Db
Qy	58	HMQOAAERKRNQYOG--SAHVSVLDD--KDLVVERFGDTNAKIE-----LPIRK 102	149	Qy
Db	150	AMARENGKKGVEENGRNPLMSALSHYAKPSLMDSEGBTWKMMLDTSPNTWDIDRNKRPPLSLPLDK	209	Db
Qy	98	--LPIRKHIDLIDVFRKKRGRMLHPKVALTIREAKSLKOLPNISPVSTAVSQQTVCGDHLGK	162	Qy
Db	210	TSLPLDKPQVAMTRAF--KVNKLHPKVLMLHEARKKIPKAMPSPVSRISTISNQTCIGDHLGK	260	Db
Qy	155	VCDLGKGLDVLVLLHKGNGLPSSNPYVNGDFDRGKGLYKULLLISLYLAFFPNVFLRNENHED	214	Qy
Db	268	ICGDLHGKFDDCILYKNGYPSVDNPYFNGDFDRGKGLYKULLLISLYLAFFPNVFLRNENHED	327	Db

QY	328 LNRGHEDSVNARYGIREVSKYPRHNRKTLADEVYRMLPPLGSVLNSRLVILWGG 274 Db	Db	449 EFRKDPKDPSGYISISHWCKMVENTUKLGPLWRLDKLAPGTDQSOKVNVRTL----- 502 QY	
QY	275 SDSTSIDLKSIDRKGVSIARPPPLTDPBLPK-----TEWOQIDIMSDPQATPGC 327 Db	Db	523 YPAKLDRTRKGSKVWKIEMTMRVNLNDLPWPLFLRGYLVADDETRVWSPHLKFHN 582 QY	
QY	388 SDQTEVKSLDKIPRHRQFVSRVPPVNGKMESENKAUNVDRKQMDIMSDPQKONGC 447 Db	Db	503 --DLDTDVTILERADGMWSMDLYANKASLVAFTNTIDDNSGRBTIDFETALIDLVA 560 QY	
QY	328 VPNTLARGAGWFGPPDTDNFLQRHSYVIRSHHECKNGHEFMHDNKITIFSAINTAI 387 Db	Db	583 FQPLMLNDWQSA---CHHTQQORANHRSQY---VAKPNKEQVSYNFCCSVIRAIDY 635 QY	
QY	448 WPNVFRGGSYFGAIDTASFLRFLKHFLLVASHRKCEPGEYFSHNNCLTVEFSAINTY 507 Db	Db	561 HMGAGAYSKAEMLKCRMGLNGOCKUDNEP-----LEAFRLSD 676 QY	
QY	388 GSNKGAYIR-LNNQLMHFVOYISAASQTKRUL-FKORMGTIVESALKELAVRMRDRE 445 Db	Db	636 TM---S2QFLPVLFVYDGGTRHIDGPKFVNMLSEMAPYPLD 676 QY	
QY	508 GSINRGAIVVKPKGSKOPHPHQYM-ASKTHRKSITLRERLGIVVEBSAVKELKEKUSSPHTD 565 Db	Db	583 FQPLMLNDWQSA---CHHTQQORANHRSQY---VAKPNKEQVSYNFCCSVIRAIDY 635 QY	
RESULT 4	446 LEDERKRYDPKDSGYISISHCKMVENTUKLGPLWRLDKLAP-GTDQSOKVNVRTL 504 Db	Db	449 EFRKDPKDPSGYISISHWCKMVENTUKLGPLWRLDKLAPGTDQSOKVNVRTL----- 502 QY	
C;Date:	566 LQKERBTMIDTRKSGKUPILKNDCVTEVIGLAWLPIAKPVATLSBEGKVKYKDRRI 625 C;Accession:	505 LTDIVILEABADGMWSMDLYANKASLVAFTNTIDDNSGRBTIDFETALIDLVA 560 C;Accession:	564 AQVGSGTHAQEID--IYESLURHKSTLTETRPFMDKNSGQVSKMSPIDACEVL----G 677 R;Anderson, B.; Allund, L.; Petterson, U.	
QY	565 AYSKARM----LEKCRMMDANGDKVLDNLFLEARLSD 599 Db	Db	626 AQVGSGTHAQEID--IYESLURHKSTLTETRPFMDKNSGQVSKMSPIDACEVL----G 677 QY	
A;Description:	93.4 kb of complete sequence from chromosome 3 of Trypanosoma cruzi.	A;Description:	565 AYSKARM----LEKCRMMDANGDKVLDNLFLEARLSD 599 A;Reference number:	678 KYTKRPLQTDYIQLAESIDFNKDGFDIDNLNELLAEFLVD 717
A;Accession:	T14614	A;Accession:	A;Accession: A55346; MUID:94357899; PMID:8077208 A;Molecule type: mRNA A;Residues: 1-499 <BEC> A;Cross-references: UNIPARC:UPI000172827; GB:X77237 A;Note: authors translated the codon AAG for residue 53 as Gln, and GTA for residue 499 C;Species: Trypanosoma cruzi C;Superfamily: protein phosphatase 5; phosphoesterase core homology; phosphoprotein phosphatase PP5 C;Keywords: iron; metalloprotein; nucleus; phosphoric monoester hydrolase; zinc C;Accession:	565 AYSKARM----LEKCRMMDANGDKVLDNLFLEARLSD 599 A;Status: preliminary; translated from GB/EMBL/DDBJ A;Molecule type: DNA A;Residues: 1-704 <END> A;Cross-references: UNIPARC:UPI000172827; EMBL:AF052832; NID:93063540; PID:93063551; PI C;Genetics: C;Map position: 3
Query Match	Best Local Similarity 17.2%; Score 593.5; DB 2; Length 704;	Query Match	16.5%; Score 569; DB 1; Length 499;	
Matches	173; Conservative 33.0%; Pred. No. 1.5e-29; Indels 75; Gaps 15;	Matches	143; Conservative 33.7%; Pred. No. 3.3e-28; Mismatches 140; Indels 68; Gaps 16;	
QY	128 ILRRAKSLKOLPN----SPVSTATS-----QQTYVCGDLHGKUDLAVLHQLNGLP 176 Db	Db	11 FICKWYRHARREMEMORRCNWQIFONLVEASEDDQAEELKKPFENDLIKMPQAERKNOQ 70 QY	
QY	174 VFTDAMSHLNTMPNTVRLSPPPVGAVRSNGRINQGSKVWWVGDHLQGLADLAHLIKCGMP 233 Db	Db	95 YIKGYYRDRASNMALGK-----FR-----ALRD-----YETVVKVKNDKDAKMQ 137 QY	
QY	177 SSSNPVVFNGDFDGRKRGLEVLILLISLYAPPNAVFLNGHEDSVNARYGIREVE 236 Db	Db	71 -----GSAH-VSLLDKD----DIVERSEG--DIVNAKELPLRKHID 107 QY	
QY	234 NEGTYTFINGDFDVRGANGEVVLLILPSMLACPKVTLNKGHECDYMDBYGFDEVN 293 Db	Db	138 ECSSKIVKQKAFERAIAQDEBHRASVVDSDIDESMTIEDEVSKPKLEGKVITFMD----- 193 QY	
QY	237 SKYPRNHKRTLAFIDBVYRMLPLGSYTLNSRVLIVHGGP--SDSTSIDLKSIDRKGVSI 294 Db	Db	108 LIDQVFKRKGKRNLRPKVYALLRBAKSLLQLENISPVSTAVSQVTWCDLHGKDDLU 167 QY	
QY	294 TKYDKN--VFLVRCFCALPLATIGKKVVFVHAGLPRRKGVNTIBDISRQFQI--348 Db	Db	194 LMKQWYKDKOK--KJURKCAIQYLQVQKEVCKLKLSTLVEVTEKIVTGDTGQDFYDIL 251 QY	
QY	295 LRPPTDGPBLDKTMRQOQF-DIMMSDPOATMGCVNTLARGAGWFGPDVTNPLQRHL 353 Db	Db	168 WVKHNGKLPSNSNPYVFGDFDVRGKRGLEVLILLISLYAPPNAVFLNGHEDSVNARYGIREVE 227 QY	
QY	349 ---PMFD--YSQPBDERDIFDPLWDMDPVBDLQHRSBPGAGVFGADVQTOFRQGLNGL 402 Db	Db	252 NTPELNGLPSSTNPYVFGDFDVRGKRGLEVLILLISLYAPPNAVFLNGHEDSVNARYGIREVE 227 QY	
QY	354 SYVRSHECKNGHESFMHDNKITTSASNTYAGNSKGAY-IRNNQLMHFVQYISA 412 Db	Db	228 RICPIREVESKYPFRNKHILAFIDBVYRMLPLGSVLSVLNSRLVILWGGP--SDSTSIDLK 285 QY	
QY	403 ELVVISHECLRGYSSHHDGKLLTIVSASNTDGPETNFGSFAVPGDNPERSFHTQVAE 462 Db	Db	312 IYCPFGEGYKAKY---ADMELSESEFWLPLAQCINGKVLIMGGLSESDGVTLDRK 368 QY	
QY	413 SQTK-----RLSKFQRMGIVTESA-----LKLAVRQEDRDRBLED 448 Db	Db	366 NFQHRHLSYVRSHECKNGHESFMHDNKITIFSAINTYAGNSKGAYIRL-NNQLMHF 404	

Db 413 AFLBENQDLYIIRSHEVKAGYEVAHGGRCVTVFSAPNYCIDOMGNKASYIHLQGSDLRPQ 472  
 Qy 405 FVQY 408  
 Db 473 HQP 476

**RESULT 6**

N;Alternate names: serine/threonine phosphatase PP5  
 C;Species: Homo sapiens (man)  
 C;Accession: S22570; PC4136  
 C;Date: 03-Aug-2001 #sequence\_revision 03-Aug-2001 #text\_change 09-Jul-2004  
 R;Chen, M.X., McPartlin, A.B.; Brown, L.; Chen, Y.H.; Barker, H.M.; Cohen, P.T.W.  
 EMBL J. 13, 4278-4290, 1994  
 A;Title: A novel human protein serine/threonine phosphatase, which possesses four tetratricopeptide repeats and a kinase domain  
 A;Reference number: S52570; MUID:95009939; PMID:7925273  
 A;Molecule type: mRNA  
 A;Residues: 7-98 <C12>  
 A;Cross-references: UNIPROT:P53041; UNIPARC:UPI0000172926; EMBL:S73586; EMBL:X09416  
 A;Experimental source: teratocarcinoma cell line NTera-2  
 R;Xu, X.-L.; Lagercrantz, J.; Zickler, P.; Balica-Lagercrantz, S.; Zetterberg, A.  
 Bloch, J.; O'Callaghan, M.; Parsons, J.; Percy, C.; Rifkin, L.; Roopra, A.; Saunders, D.  
 A;Reference number: PC4136; MUID:9514708; PMID:8561788  
 A;Accession: PC4136  
 A;Residue type: RNA  
 A;Cross-references: UNIPARC:UPI000016B012; EMBL:X92121; NID:9117477; PIDN:CAA63089.1;  
 A;Experimental source: fetal brain  
 C;Comment: mRNA encoding this protein was detected in all human tissues tested and was  
 C;Genetics:  
 A;Gene: GDB:PPSC; PPPS; PP5  
 A;Cross-references: GDB:136857  
 A;Map position: 19q13.3-19q13.3  
 C;Function:  
 A;Description: catalyzes the hydrolytic dephosphorylation of protein-phosphoserine [vali  
 A;Note: may play a role in the regulation of RNA synthesis and mitosis  
 C;Superfamily: protein phosphatase 5; phosphoesterase core homology; phosphoprotein phos  
 C;Keywords: iron; metalloprotein; nucleus; phosphoric monoester hydrolase; zinc  
 F;2861/Domain: tetratricopeptide repeat homology <TT1>  
 F;62-95/Domain: tetratricopeptide repeat homology <TT2>  
 F;96-129/Domain: tetratricopeptide repeat homology <TT3>  
 F;204-67/Domain: phosphoprotein phosphatase homology <PPP>  
 F;236-305/Domain: phosphoesterase core homology <BBC>  
 F;244-271/Binding site: iron (Asp, His, Asp) #status predicted  
 F;271, 303-352, 426/Binding site: zinc (Asp, Asn, His, His) #status predicted  
 F;274, 304-450/Active site: Asp, His, Tyr #status predicted  
 F;275, 399/Binding site: substrate phosphate (Arg) #status predicted

Query Match 16.5%; Score 46.5; DB 1; Length 498;  
 Best Local Similarity 34.0%; Pred. No. 4.1e-88; Matches 144; Conservative 70; Mismatches 142; Indels 67; Gaps 16;

Qy 11 FIQKWRHRQAERREMORCNMQIFONLEYASQDQARLYKPFNDLKHMPOAAGRNQO 70  
 Db 95 YIKGYYRRAANMAGK-----FR----ALARD-----YETVVKVPKHDQAKMYQ 137  
 Qy 71 -----GSAHV-SVLDDKD---DLVERFG--DIVAKIEPIRKHIDL 107  
 Db 138 ECKIVKQKAFFERAIADEBHKSVPSLSDISWTSIDBEYSQPKLEGKVTSFMSKE--- 193  
 Qy 108 LDIVFRKKRKGRLHPKVVALIREEAKSLQKLPNISPVASTAVSQVTVCGDLHKGKUDLL 167  
 Db 194 IAKWYKQOK-KLHKCKAQYQJUVQVKVLSKLSTLVETTAKTEKTVCGDTHQFDLL 251  
 Qy 168 VVLAHQGLPSSNPYVNGDFDRKGKLGELVILLASLYLAPPNAVFLNRGRHEDVMNA 227  
 Db 252 NIFELNGLPSETNPYIENGDFVDRGSFSVEVLTLEFGFKLXPDHFILRGRHETDNMQ 311  
 C;Or 228 RVGPIREVESKVPRMHCRILAFIDEVYMLPLGSVINSRLIVHGGPSDST-SLDIKSI 286

**RESULT 7**

N;Alternate names: serine/threonine phosphatase (EC 3.1.3.16) Y99B6B\_ff [similarity] - Caenorhabditis elegans  
 C;Species: Caenorhabditis elegans  
 C;Accession: T45058  
 C;Date: 03-Aug-2001 #sequence\_revision 03-Aug-2001 #text\_change 28-Apr-2003  
 C;Accession: T45058  
 R;Wilson, R.; Aliscough, R.; Anderson, K.; Baynes, C.; Beake, M.; Bonfield, J.; Burton, R.; Furlton, I.; Gardner, A.; Green, P.; Hawkins, T.; Hillier, L.; Jier, M.; Johns, B.; O'Callaghan, M.; Parsons, J.; Percy, C.; Rifkin, L.; Roopra, A.; Saunders, D.; Nature 368, 32-38, 1994  
 A;Authors: Shawoneen, R.; Sims, M.; Smalton, N.; Smith, A.; Smith, M.; Sonnhammer, E.; Stoeck, L.; Wilkinson-Sproat, J.; Wohldman, P.  
 A;Title: 2.2 Mb of contiguous nucleotide sequence from chromosome III of C. elegans  
 A;Reference number: 943531; MUID:94150718; PMID:7906398  
 A;Accession: T45058  
 A;Status: translated from GB/EMBL/DDJB  
 A;Molecule type: DNA  
 A;Residues: 1-526 <WIL>  
 A;Cross-references: UNIPARC:UPI0000083436; EMBL:AL132896; NID:95434440; PIDN:CAA60937.1;  
 A;Experimental source: clone Y39B6B  
 C;Genetics:  
 A;Map position: 3  
 A;Introns: 42/1; 124/3; 184/3; 265/3; 348/3; 385/2; 415/2; 499/3  
 A;Note: Y39B6B\_ff  
 C;Superfamily: protein phosphatase 5; phosphoesterase core homology; phosphoprotein phos  
 C;Keywords: iron; metalloprotein; nucleus; phosphoric monoester hydrolase; zinc  
 F;29-61/Domain: tetratricopeptide repeat homology <TT1>  
 F;62-95/Domain: tetratricopeptide repeat homology <TT2>  
 F;96-129/Domain: tetratricopeptide repeat homology <TT3>  
 F;203-497/Domain: phosphoprotein phosphatase homology <PPP>  
 F;235-304/Domain: phosphoesterase core homology <BBC>  
 F;241, 243-270/Binding site: iron (Asp, His, Asp) #status predicted  
 F;270, 302, 351, 456/Binding site: zinc (Asp, Asn, His, His) #status predicted  
 F;273, 303, 420/Active site: Asp, His, Tyr #status predicted  
 F;274, 429/Binding site: substrate phosphate (Arg) #status predicted

Query Match 16.2%; Score 538; DB 1; Length 526;  
 Best Local Similarity 31.6%; Pred. No. 1.7e-27; Matches 138; Conservative 79; Mismatches 172; Indels 48; Gaps 14;

Qy 3 ENAIRAI---FIQKWRHRQAERREMORCNMQIFONLEYASQDQARLYKPFNDLKHMPOAAGRNQO 56  
 Db 86 DMAL---AIDSYKVKPRTTANMALLGRKALTYQAVVTKCUPDKRAXFKFCSKIV 143  
 Qy 57 KHMPOAAGRNQOYQSAHVSLDDKDVLVERFGTVNKAKTEPLTKHIDLIVDPRKK 116  
 Db 144 RROKKEBAISTDHDKKTVAEIDINAMAIKEDSYD-GPRLIEDKTKFVQLIKTFRNQO 201  
 Qy 117 GNLRKPKVVALIREEAKSLQKLPNISPVASTAVSQVTVCGDLHKGKUDLL 176  
 Db 202 --KLHKCKAQYQJUVQVKLSTLVETTAKTEKTVCGDTHQFDLL 259  
 Qy 177 SSSNPYVNGDFDRKGKLGELVILLASLYLAPPNAVFLNRGRHEDVMNA 236  
 Db 260 SETNPYIENGDFVDRGSFSVEVLTLEFGFKLXPDHFILRGRHETDNMQ 319

QY 237 SKYPRNNHKKRLATIDDVIRWLPLOSSVLSRVLVHGGF--SDSTSIDLKSIDRGKVSI 294  
 C;Species: Neurospora crassa  
 C;Accession: 03-Aug-2001 #sequence\_revision 03-Aug-2001 #text\_change 28-Apr-2003  
 C;Status: translated from GB/EMBL/DDBJ  
 R;Yatzkan, B.; Yarden, O. Library, February 1997  
 submitted to the EMBL Data Library, February 1997  
 A;Description: Ppt-1, a N. crassa novel-type phosphatase  
 A;Reference number: Z23069  
 A;Accession: T46576  
 A;State: translated from GB/EMBL/DDBJ  
 A;Molecule type: mRNA  
 A;Residue: 1-479 <YAT>  
 A;Cross-references: UNIPARC:UPI00000D260; EMBL:U09985; PIDN:AAB65138.1  
 C;Genetic:  
 A;Gene: ppt-1  
 A;MP Position: V  
 C;Superfamily: protein phosphatase 5; phosphoesterase core homology; phosphoprotein phosphatase  
 C;Keywords: iron; metalloprotein; nucleus; phosphoric monoester hydrolase; zinc  
 P;8-41/Domain: tetratricopeptide repeat homology <TT1>  
 P;8-42/75/Domain: tetratricopeptide repeat homology <TT2>  
 P;76-109/Domain: tetratricopeptide repeat homology <TT3>  
 P;103-149/Domain: phosphoprotein phosphatase homology <PPP>  
 P;215-284/Domain: phosphoesterase core homology <BBC>  
 P;211-223, 250/Binding site: iron (Asp, His, Asp) #status predicted  
 P;220, 282, 331, 408/Binding site: zinc (Asp, His, Asn, His, His) #status predicted  
 P;233, 283, 432/Active site: Asp, His, Tyr #status predicted  
 P;254, 381/Binding site: substrate phosphate (Arg) #status predicted  
 Query Match 15.7%; Score 541; DB 1; Length 479;  
 Best Local Similarity 32.3%; Pred. No. 1.8e-26;  
 Matches 140; Conservative 57; Mismatches 149; Indels 88; Gaps 10;  
 QY 71 GSARHVSVDQD---DLVR-----EGDIYNAKTELPPIRKNIHLIDP---110  
 Db 108 -----DNDAKLUKIVCEKRVQLRFAIEVDELSAEGLDVESMAVDAISDGVR 159  
 QY 111 -----VFRKKGRGRNLHKKVVALILREAKSLKLQLNPNSPVSTAVSQ 156  
 Db 75 FVKAYVRRATA-----YAAILNPKEAVKDKPKTCVKLAP-----107  
 QY 71 GSARHVSVDQD---DLVR-----EGDIYNAKTELPPIRKNIHLIDP---110  
 Db 108 -----DNDAKLUKIVCEKRVQLRFAIEVDELSAEGLDVESMAVDAISDGVR 159  
 QY 111 -----VFRKKGRGRNLHKKVVALILREAKSLKLQLNPNSPVSTAVSQ 156  
 Db 75 FVKAYVRRATA-----YAAILNPKEAVKDKPKTCVKLAP-----107  
 QY 94 AKLR-LPIRKNIHLILVFRKGRGRNLHKKVVALILREAKSLKLQLNPNSPVSTAVSQ 152  
 Db 214 ARIEGBVFLDFVTKMMEDPKNGK--TLMKRYAVOVIQLOTRQIILALPSLVIDSVPHGRH 271  
 QY 153 VTVCGDGLGKDDLLVVVLEKNGLSSSNPYVFGNDFVDRGKGGLFLVLLLSSLYLAEPNA 212  
 Db 272 ITVCGDGLGQFYDPLNIFELANGPSPBENYLNGDFVDRGSPSVEILITFARKCMCFS 331  
 QY 213 VFLNRGNHEDSVNARYGTFIREVSKYPRHKRKLAPIID--EVYRMLLGSVNRSVIL 269  
 Db 332 IYIARGNHESKSMMKIKYGEGERTRKLSER----FVDLPAEKFVCPFLAHVINGKVRF 385  
 QY 270 VHGG-FS-DSTSIDLKISIDRGKXVISLRLPPLDGEPLDKTEWQOIDMWSDQATMCVPNTL 327  
 Db 386 VHGG-FS-VGGLSVTGVKULSDTRAIDR----FCEP-----EBCIMCBLLWDSPOLPER 430  
 QY 328 VPTVLRLGAGWFGSPVTDMFLQRHLSVIRSHECKPKNGHEFMENDNKLTITFASNYTAI 387  
 Db 431 GPSK-RGVLGFSGFGDVTFRQDQINLILUVRSHEVKDGSYEVSHDGKLITVFSAPNYCDQ 489  
 QY 388 GSNKGVAYIRLN-NQLMPHFVQY 408  
 Db 490 MGKNGAFIRFEADMKPNIVTP 511

RESULT 9

T46576 Phosphoprotein phosphatase (EC 3.1.3.16) Ppt-1 [similarity] - Neurospora crassa  
 N;Alternate names: Serine/threonine phosphatase Ppt5 homolog  
 C;Species: Schizosaccharomyces pombe  
 C;Date: 03-Aug-2001 #sequence\_revision 03-Aug-2001 #text\_change 28-Apr-2003  
 C;Accession: T40391  
 R;Lyne, M.; Randalream, M.A.; Barrell, B.G.; Churcher, C.M.  
 submitted to the EMBL Data Library, February 1998

RESULT 10

T40391 Phosphoprotein phosphatase (EC 3.1.3.16) SPBC3P6.01C [similarity] - fission yeast (Schizosaccharomyces pombe)  
 N;Alternate names: Serine/threonine phosphatase Ppt5 homolog  
 C;Species: Schizosaccharomyces pombe  
 C;Date: 03-Aug-2001 #sequence\_revision 03-Aug-2001 #text\_change 28-Apr-2003  
 C;Accession: T40391  
 R;Lyne, M.; Randalream, M.A.; Barrell, B.G.; Churcher, C.M.  
 submitted to the EMBL Data Library, February 1998

A;Reference number: Z21925  
A;Accession: T40391  
A;Status: translated from GB/EMBL/DBBJ  
A;Molecule type: DNA  
A;Residues: 1-473 <LNN>  
A;Cross-references: UNIPARC:UPI000006B641; EMBL:AL022019; PTDN:CAA17690\_2; GSPDB:GN0007  
A;Experimental source: strain 972h-, cosmid c3F6  
C;Genetics:  
A;Map position: 2  
C;Superfamily: protein phosphatase 5; phosphoesterase core homology; phosphoprotein phosphatase; zinc finger repeat homology <TT1>  
F;39--/2/Domain: tetrastricopeptide repeat homology <TT2>  
F;73--10/2/Domain: tetrastricopeptide repeat homology <TT3>  
F;179--44/2/Domain: phosphoprotein phosphatase homology <PPP>  
F;211--280/2/Domain: phosphoesterase core homology <PPC>  
F;217--219/2/Domain: binding site: iron (Asp, His, Asp) #status predicted  
F;246--278--327/40/2/Domain: binding site: zinc (Asp, Asn, His, His) #status predicted  
F;249--279--428/2/Domain: active site: Asp, His, Tyr #status predicted  
F;250--377/2/Domain: binding site: substrate phosphate (Arg) #status predicted

Query Match Best Local Similarity 35.1%; Score 519; DB 1; Length 473; Matches 119; Conservative 66; Mismatches 127; Indels 26; Gaps 8; C;Keywords: iron; metalloprotein; nucleus; phosphoric monoester hydrolase; zinc

Qy 73 ARVSVLDDKDDLVTEEGDIDVNAKEPLPIRKONHDLIDVPRKRCRGRNLHPPKVALLREA 132  
Db 139 ANINIED--MDPSDYDCVII--LEKQITKEFVEDMKRF--CQGKKLPLPKAYSLRDL 191  
Qy 133 AKSLQQLPNTISPVSTAVSQQVTYCGDLHKGKDDLLVILVHKNGLPSSSNPNTYFGDNDVRG 192  
Db 192 KELLEKTPPSLIDIPVKDGTETLVCIGDPTGQYDPLNFITKLKGPPSPTPNPKVFLNGDVRG 251  
Qy 193 KRGLEVLLLISLTYLAFFNAVFLNRGHEDDSMMARYGFIREVESKYPKRHNRLAPIDE 252  
Db 252 SMSTEVARTFLYAVKGLYDPAVFINRGHEDDMMARVGFSEGRCSRKY--WERTFNPSE 308  
Qy 253 VVWMLPLGSVLMRSARVLVTHGGF--SDS7SLDLIKTSDRGKVSILRPLTGPBRPLDKTEW 310  
Db 309 TFSLPLPLGSLSIDSYLVVHGGLRSDDNDVTLQDOLRNIDRFS---KQPGQSG----- 356  
Qy 311 QDQFDIMWSDPOATMGCVNPNTURGAGWGMGPPYDNTLQRHLSVTRSHCKPNHFRM 370  
Db 357 :-----LMMELWLTDPQAPGRGPK-RGVGLQGPQDFSKRCFCEANGLAKAVIRSHVRDQCYVE 414  
Qy 371 HDNKLIITPSASVYIALGSNGKAYIRLNQJLMPHFVQ 408  
Db 415 HOGYCITYFSAPYCDSTGNTLGAIVKVKYKBDMSLDHQFQ 452

RESULT 11

S52571 Phosphoprotein phosphatase (EC 3.1.3.16) PP1 - yeast (Saccharomyces cerevisiae)  
C;Species: Saccharomyces cerevisiae  
C;Date: 03-Aug-2001 #sequence revision 03-Aug-2001 #text\_change 09-Jul-2004  
C;Accession: S52571; S55981; S64432; S64697  
R;Chen, M.X.; McPartlin, A.E.; Brown, L.; Chen, Y.H.; Barker, H.M.; Cohen, P.T.W.  
R;MOU\_J: 13\_4274-4290\_1994  
A;Title: A novel human protein serine/threonine phosphatase, which possesses four tetrastricopeptide repeats, remnants of Ty and three tRNA genes.  
A;Reference number: S52570; MUID:95009929; PMID:7925273  
A;Accession: S52571  
A;Molecule type: mRNA  
A;Residues: 1-1513 <CH>  
A;Cross-references: UNIPROT:P53043; UNIPARC:UPI0000053298  
R;van Dyck, L.; Goffeau, A.  
A;Submitted to the EMBL Data Library, December 1994  
A;Description: Genes for an asn synthase, a G1P-motif nucleoporin and a putative homeobox gene ORFs, remnants of Ty and three tRNA genes.  
A;Reference number: S55976  
A;Accession: S55981  
A;Molecule type: DNA

RESULT 12

T51611 Phosphoprotein phosphatase (EC 3.1.3.16) PP7 [imported] - Arabidopsis thaliana  
C;Species: Arabidopsis thaliana (mouse-ear cress)  
C;Date: 18-Aug-2000 #sequence\_revision 18-Aug-2000 #text\_change 09-Jul-2004  
C;Accession: T51611

A;Residues: 1-513 <WAN>  
A;Cross-references: UNIPARC:UPI0000053298; EMBL:X83099; NID:9642340; PID:9642346  
A;Experimental Source: strain S28C  
R;Chen, M.X.; McPartlin, A.E.; Brown, L.; Chen, Y.H.; Barker, H.M.; Cohen, P.T.W.  
A;Submitted to the EMBL Data Library, July 1995  
A;Description: A novel human protein serine/threonine phosphatase, which possesses four tetrastricopeptide repeats, which is similar to the UNIPARC entry S52571.  
A;Accession: S64497  
A;Molecule type: DNA  
A;Residues: 1-380, 'HHL', 389, 'AP', 392-393, 'NFGYGRTHKPKTGWL', 409, 'SVV' <CHW>  
A;Cross-references: UNIPARC:UPI00016881; EMBL:X89417; NID:987805; PID:987806  
C;Genetics:  
A;Map position: 7R  
C;Superfamily: protein phosphatase 5; phosphoesterase core homology; phosphoprotein phosphatase; zinc finger repeat homology <TT1>  
F;46--79/2/Domain: tetrastricopeptide repeat homology <TT2>  
F;80-113/2/Domain: tetrastricopeptide repeat homology <TT3>  
F;209-475/2/Domain: phosphoprotein phosphatase homology <PPP>  
F;243--312/2/Domain: phosphoesterase core homology <PPC>  
F;249, 251-278/2/Domain: binding site: iron (Asp, His, Asp) #status predicted  
F;278, 310-335, 434/2/Domain: binding site: zinc (Asp, Asn, His) #status predicted  
F;281, 311, 458/2/Domain: active site: Asp, His, Tyr #status predicted  
F;282, 407/2/Domain: binding site: substrate phosphate (Arg) #status predicted

Query Match Best Local Similarity 32.3%; Score 479.5; DB 1; Length 513; Matches 130; Conservative 68; Mismatches 154; Indels 51; Gaps 13; C;Keywords: iron; metalloprotein; nucleus; phosphoric monoester hydrolase; zinc

Qy 2 DENATAIAPIQKWMYRHRQARRE--ORQCNWQIQLQYVASEBQDQABLYKFENDLKH 58  
Db 113 DPAATKALKTCDFRIFERFRKAIGGAENAEAKISLQCTQNLSSFPDNADLANY----- 165  
Qy 59 MPOAQAKRKNQYQGSASVSVLDDKODIVDVAKEPLPIRKONHDLIDVFRKGRN 118  
Db 166 -----SOPKLSPE-----QLYDQN-----AFKQAKKIKSOMQBTISKMNDLTK-GK 207  
Qy 119 RLHPKTYVALTIREAAKSQKOLPNTISPV--STAVSQOVTGCGDLHKGKDDLLVILVHKNGLP 176  
Db 208 YLPKKVVAIISHADTFIQPSWMELENNSTPDVKISVCGDTHQFYDVNLNFKFGKV 267  
Qy 177 SSSNPVTFNGDPYDVGKRGLEYLULSLYLAFPNAVFLPGLRGNHSDSMMARYGIRRE 236  
Db 268 GPKHTVLFNQDPVDRGSSWSCEVALLYFLCKLTKLHPNNPFLPQNGHNSDDNNMKVYGFEDCK 327  
Qy 237 SKYPRHKKRILAPIDEVRYMLPLGSVLSNRLVHGG-SD-STSLDLKSIDRKGKVI 294  
Db 328 YKYY--S-QRINTMPPASQESPLATLNUYLUWIGGLPDPSPATLDFKNIDR----F 379  
Qy 295 LRPLPTEGEPLDKTEWQDIDMWSDDPOATMGCVNTLRGAGWMPDQDVDFNLFHRLS 354  
Db 380 AQPP--RDG-----AFMELLWADPQEANGMGSQ-RLGLHAFQDIDTRFLRANKLR 428  
Qy 355 YVTRSHCKPNHGRFMHDNKKITISASNTYAGENKGAYIRL 397  
Db 429 KLFRSHELRMGGQVFBQGKLMUTVFSAPNYCDSQGNGLGIVHV 471



RESULT 16

Best Local Similarity 30.7%; pred. No. 6e-20; Matches 102; Conservative 65; Mismatches 128; Indels 37; Gaps 7



RESULT 20

Db 88 LFEYGGPPBEAN-YLFGLDYYDRGKOSLETICLLAYKIKYPENFLRGNHESASINRI 146  
 Qy 229 YGPIRKVESKVPRNHKILAFIDETYRMLPLGSVLNSRVLVHGAFS-DTSLDLKSID 287  
 C;Species: Arabidopsis thaliana (mouse-ear cress)  
 C;Date: 16-Sep-1992 #sequence\_revision 16-Feb-2001 #text\_change 15-Mar-2004  
 C;Accession: S20882; S31085; S49037; H8495  
 R:Nitschke, K.; Pleig, U.; Scheil, J.; Palme, K.  
 EMBL J. 11, 1327-1333, 1992  
 A;Title: Complementation of the cs dis2-11 cell cycle mutant of *Schizosaccharomyces pombe*  
 A;Reference number: S20882; MUID:92224871; PMID:1314161  
 A;Status: nucleic acid sequence not shown  
 A;Residues: 1-1318 <NTT>  
 A;Cross-references: UNIPARC:UPI0000175977; EMBL:X64328; NID:916430; PIDN:CAA45611.1; PII  
 R:Smith, R.D.; Walker, J.C.  
 Plant Mol. Biol. 21, 307-316, 1993  
 A;Title: Expression of multiple type 1 phosphoprotein phosphatases in *Arabidopsis thaliana*  
 A;Reference number: S31085; MUID:93144705; PMID:7678768  
 A;Accession: S31085  
 A;Status: nucleic acid sequence not shown  
 A;Molecule type: mRNA  
 A;Residues: 1-1318 <SMI2>  
 A;Cross-references: UNIPARC:UPI0000175977; EMBL:M93408; NID:9166571; PIDN:AA32723.1; PI  
 R:Lin, X.; Kaul, S.; Rounseley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;  
 M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umeyam, I.; Tallon, L.;  
 Eubis, D.; Niernberg, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.  
 Nature 402, 761-768, 1999  
 A;Title: Sequence and analysis of chromosome 2 of the plant *Arabidopsis thaliana*.  
 A;Reference number: A84420; MUID:20083487; PMID:10517197  
 A;Accession: H84695  
 A;Status: preliminary  
 A;Residues: 1-1318 <STO>  
 A;Cross-references: UNIPARC:UPI0000175977; GB:AE002093; NID:93980395; PIDN: AAC95198.1.; G  
 C;Genetics:  
 A;Gene: At2g29400  
 A;Map position: 2  
 C;Superfamily: serine/threonine protein phosphatase; phosphoesterase core homology; phosphoesterase; experimental source; clone TOPPI  
 C;Keywords: iron; metalloprotein; phosphoric monoester hydrolase; serine/threonine-specific phosphatase; F41-300/Domain: phosphoprotein phosphatase homology <PPP>  
 F69-137/Domain: phosphoesterase core homology <PEBC>  
 Query Match 12.4%; Score 426.5; DB 2; Length 317;  
 Best Local Similarity 30.2%; Pred. No. 1.6e-19;  
 Matches 106; Conservative 62; Mismatches 132; Indels 51; Gaps 8;  
 A;Status: preliminary  
 A;Residues: 1-1318 <STO>  
 A;Cross-references: UNIPARC:UPI0000175977; GB:AE002093; NID:93980395; PIDN: AAC95198.1.; G  
 C;Genetics:  
 A;Gene: At2g29400  
 C;Superfamily: serine/threonine protein phosphatase; phosphoesterase core homology; phosphoesterase; experimental source; clone TOPPI  
 C;Keywords: iron; metalloprotein; phosphoric monoester hydrolase; serine/threonine-specific phosphatase; F41-300/Domain: phosphoprotein phosphatase homology <PPC>  
 F71-139/Domain: phosphoesterase core homology <PEBC>  
 F77-79, 105/Binding site: iron (Asp, His, Asp) #status predicted  
 F105, 137-186, 226/Binding site: zinc (Asp, Asn, His, His) #status predicted  
 F109, 233/Binding site: substrate phosphate (Arg) #status predicted  
 Query Match 12.4%; Score 428; DB 2; Length 318;  
 Best Local Similarity 29.6%; Pred. No. 1.3e-19;  
 Matches 106; Conservative 67; Mismatches 121; Indels 64; Gaps 11;  
 Db 17 117 -GHLILHPKVVALLIREAKSLKLQPNLTS-----PVSTAVSQVTWCGDLGKQDLY 168  
 Db 37 SGKQVH-----LSEG-BIRQLCAVSKEIFLQQPNLLEAPIKICGDIQGQSDILR 87  
 Qy 59 MPQAAGRENOQYQSAHVSVTDKDDUYVERGDIVNAKTEPLIRKRNHIDLUDVPRKKR- 116  
 Db 1 1 MARKPAQDQSKQRKMRAPVLD-----DIRRAVE-----FRVTRPG 36  
 Qy 203 -----PTDVPDSGLCDLIMSDPSREVKGWGMNDRGKSYTFQDPKVAEFLMHD 253  
 Db 288 RGYVSLTRPLTDGCBPLDKYEWQOIFFDIMSMDPQATMGCVPTLRRGAGWFGPDPDVNFQHRRLS 354  
 Db 204 R-----PMIDPESGLVQCDLWMSDPGSDVGNG-MND-RGVSTFQADKVAEF 247  
 Qy 348 LQRHRLSVIYRSHECKNGRFHMHNKIKITFSASNYTAIGSNKARVIRUNNQMPHF 405  
 Db 248 LRKHDMDLICRAHQVVEDGYRFPAEROLWTVPSAPNVCGBFDNAGAMMSIDESLMCSF 305

RESULT 21

Db 88 LFEYGGPPBEAN-YLFGLDYYDRGKOSLETICLLAYKIKYPENFLRGNHESASINRI 146  
 Qy 229 YGPIRKVESKVPRNHKILAFIDETYRMLPLGSVLNSRVLVHGAFS-DTSLDLKSID 287  
 C;Species: Nicotiana tabacum (common tobacco)  
 C;Date: 24-Mar-1999 #sequence\_revision 24-Mar-1999 #text\_change 09-Jul-2004  
 C;Accession: T03594  
 R:Suh, M.; Cho, H.; Kim, Y.; Liu, J.; Lee, H.  
 Plant Mol. Biol. 36, 311-322, 1998  
 A;Title: Multiple genes encoding serine/threonine protein phosphatases and their differences  
 A;Reference number: Z14967; MUID:9814537; PMID:9484443  
 A;Status: Preliminary  
 A;Molecule type: mRNA  
 A;Residues: 1-317 <SUH>  
 A;Cross-references: UNIPROT:004856; UNIPARC:UPI0000131FAC; EMBL:Z93768; NID:9129403; PI  
 A;Experimental source: cultivar Xanthi  
 C;Genetics:  
 A;Gene: np\_1  
 C;Superfamily: serine/threonine protein phosphatase; phosphoesterase core homology; phosphoesterase; experimental source; clone TOPPI  
 F41-300/Domain: phosphoprotein phosphatase homology <PPP>  
 F69-137/Domain: phosphoesterase core homology <PEBC>  
 Query Match 12.4%; Score 426.5; DB 2; Length 317;  
 Best Local Similarity 30.2%; Pred. No. 1.6e-19;  
 Matches 106; Conservative 62; Mismatches 132; Indels 51; Gaps 8;  
 Db 59 MPOAAGRKNQYQGSARVSVAADDKODIVLVEFGDIVAKIENPIRKNHIDLUDVFRKKRN 118  
 Qy 1 MAQNHHQQQQQGQLRAGVLD-----DINRLR-----FRNARTV 36  
 Db 119 R---LHKYVALLILRAKSIKLPNISPYSTAVSQVTWCGDLGKQDLYLVWLRKGL 175  
 C;Genetics:  
 A;Gene: At2g29400  
 C;Superfamily: serine/threonine protein phosphatase; phosphoesterase core homology; phosphoesterase; experimental source; clone TOPPI  
 C;Keywords: iron; metalloprotein; phosphoric monoester hydrolase; serine/threonine-specific phosphatase; F41-300/Domain: phosphoprotein phosphatase homology <PPC>  
 F71-139/Domain: phosphoesterase core homology <PEBC>  
 F77-79, 105/Binding site: iron (Asp, His, Asp) #status predicted  
 F105, 137-186, 226/Binding site: zinc (Asp, Asn, His, His) #status predicted  
 F109, 233/Binding site: substrate phosphate (Arg) #status predicted  
 Query Match 12.4%; Score 428; DB 2; Length 318;  
 Best Local Similarity 29.6%; Pred. No. 1.3e-19;  
 Matches 106; Conservative 67; Mismatches 121; Indels 64; Gaps 11;  
 Db 176 PSSSNVYQNPDPVPGKRGKLGVLVLLSLYLAPPNAELNRGHNEDSVMARYGPIRE 235  
 Db 93 PPEAN-YLFLGDYVDRGKOSLSTICLILAYKIKYPENPFLLRGHNCASINRKYDFC 151  
 Qy 236 ESKYPRHKRILAFIDEVYRMLPLGSVLNSRVLVHGAFS-DTSLDLKSIDRGKVS 294  
 Db 152 KRPF--NVKLMKCFTCFLPVAALDEKLCMGLSPVLTMDQIRNLPR----- 202  
 Db 295 LRPPLTDGEPLDKTSMQIFDINWSDPQATMGCVPTLRRGAGWFGPDPDVNFQHRRLS 354  
 Db 203 -----PTDVPDSGLCDLIMSDPSREVKGWGMNDRGKSYTFQDPKVAEFLMHD 253  
 Qy 355 YVIRSHCKENGHRFMNDKITTISASNTAIGSNKAYTRNNQMPHF 405  
 Db 254 LVCRAHQVVEDGYRFPAEROLWTVPSAPNVCGBFDNAGAMMSIDESLMCSF 304

RESULT 22

T21288 Phosphoprotein phosphatase (EC 3.1.3.16) F23B12.1 [similarity] - *Caenorhabditis elegans*  
 C;Species: *Caenorhabditis elegans*  
 C;Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 05-Oct-2004  
 C;Accession: T21288



RESULT 25

Q9548 phosphoprotein phosphatase (EC 3.1.3.16) 1 catalytic chain delta - alfalfa  
 C;Species: *Medicago sativa* (alfalfa)  
 C;Accession: T09548  
 C;Description: sequence\_revision 16-Jul-1999 #text\_change 09-Jul-2004  
 R;Visst, B.; Csordas Toth, B.; Kovacs, I.; Magyar, Z.; Horvath, G.; Bagossi, P.; Gergely Arch. Biochem. Biophys. 360, 206-214, 1998  
 A;Title: Protein phosphatase 1 catalytic subunit isoforms from alfalfa: Biochemical characterisation and expression analysis  
 A;Reference number: Z16730; MUID:99068922; PMID:9851832  
 A;Accession: T09548  
 A;Status: preliminary; translated from GB/EMBL/DDJB  
 A;Molecule type: mRNA  
 A;Residues: 1-326 <WIS>  
 A;Cross-references: UNIPROT:O65846; UNIPARC:UPI00000AAB92; EMBL:AJ002487; NID:93176075;  
 A;Experimental source: subspecies sativa, strain RA3  
 C;Genetics:  
 C;Function: PPI delta  
 C;Description: catalyzes hydrolysis of peptidyl-phosphoserine or -phosphothreonine to release phosphate  
 C;Keywords: serine/threonine protein phosphatase; phosphoesterase core homology; phosphotyrosine phosphatase homology <PPP>  
 F;54-122/Domain: phosphoprotein phosphatase core homology <PEC>  
 Query Match 12.2%; Score 421; DB 2; Length 317;  
 Best Local Similarity 31.6%; Pred. No. 3.8e-19; DB 2; Length 317;  
 Matches 103; Conservative 61; Mismatches 123; Indels 39; Gaps 10;  
 Q9548 108 LDYFRKKGMLHPKVALLIREAKSLKOLPNIS-----PVSTAVSQCVTGCGLH 160  
 Db 22 LIEVRTARPGKQHV-----LSEAS-EIQCICSVSREBILFSPNLLERAPKIKCGDIH 72  
 Q9548 161 GKLDDLIVVHLNGLSSSSNPYVFNEDPVDGRGKGLVEVLLSLVLAFFPNVFLNRGNH 220  
 Db 73 GSYSDLRLRPFYGGPPEAN YLFLDYVDRGKQSATICLIGYKIKYPNPFLLRGNH 131  
 Q9548 221 EHSMVARYGFREVEVKYPRMKRILAFIDEVYRMPPLASVLSRVLIVGGS-DSTS 279  
 Db 132 EASINRYIYGFDCKRPF--NMRLMKVFTDCFCNLCPVAAIIDKKLICMGGLSDLTH 188  
 Q9548 280 LDLIKSDRGKGVSLRPLPTDGPRLDKTEMQIDIMSPD-QATMGCVNTLRAAGW 338  
 Db 10 INRLLBV-RGRPGKQVQ-----LSEA--EIQCICLVSKDIFMNQPNLKLKAPKIKCG 59  
 Q9548 158 DLKQKDLILVWHLNGLCPSSSPYVFGDFDVRGKGLVEVLLSLVLAFFPNVFLNR 217  
 Db 60 DIHQQYSDLRLRFYGGPPEAN YLFQDYVDRGKQSATICLIGYKIKYPNPFLLR 118  
 Q9548 218 GNHEDSVMARYGFREVEVKYPRMKRILAFIDEVYRMPPLASVLSRVLIVGGS-D 276  
 Db 119 GNHECASINRYIYGFDCKRPF--NMVLKMTDCFCNLCPVAAIIDKKLICMGGLSPE 175  
 Q9548 277 STSLDLIKSDRGKGVSLRPLPTDGPRLDKTEMQIDIMSPD-QATMGCVNTLRAAGW 336  
 Db 176 LHDNQIQNLR-----PCEVPVSGLQDLMNSDPSDSSDVRGMGERSERGS 220  
 Q9548 337 VMECPDVDTNFLQRHLSVIRSHCKPENGHERMDNKITTSASNYTAIGSNKGAYIR 396  
 Db 221 YTGFADRVKEFLQGDLDLICRAHODVGYEFPAQNQVLTFSAPNYCGERFDNAGAMVTD 280  
 Q9548 397 LNNOQMPHVOYISAASOTKRSFKRQMGIVVERSALK 433  
 Db 281 VNEELIVCSR-OQIKELDKKEPKFSGSTTVKSSPTK 316  
 RESULT 26  
 probable phosphoprotein phosphatase (EC 3.1.3.16) 1 catalytic chain - rice  
 C;Species: *Oryza sativa* (rice)  
 T03304  
 C;Accession: T03304  
 C;Description: sequence\_revision 16-Jul-1999 #text\_change 09-Jul-2004  
 R;Wang, B.; Chang, M.; Chen, X.; Wu, R.  
 submitted to the EMBL Data Library, July 1995  
 A;Reference number: Z14900  
 A;Accession: T03304  
 A;Status: preliminary; translated from GB/EMBL/DDJB  
 A;Molecule type: mRNA  
 A;Residues: 1-317 <WAN>  
 A;Cross-references: UNIPROT:P48489; UNIPARC:UPI0000131FBF; EMBL:U31773; NID:951335; PID:951335  
 A;Experimental source: seed  
 C;Superfamily: serine/threonine protein phosphatase; phosphoesterase core homology; phosphotyrosine phosphatase homology <PPP>  
 C;Keywords: phosphoric monoester hydrolase  
 F;36-395/Domain: phosphoprotein phosphatase homology <PPP>  
 F;64-132/Domain: phosphoesterase core homology <PEC>  
 Query Match 12.2%; Score 420.5; DB 2; Length 317;  
 Best Local Similarity 31.6%; Pred. No. 3.8e-19; DB 2; Length 317;  
 Matches 103; Conservative 61; Mismatches 123; Indels 39; Gaps 10;  
 Q9548 108 LDYFRKKGMLHPKVALLIREAKSLKOLPNIS-----PVSTAVSQCVTGCGLH 160  
 Db 22 LIEVRTARPGKQHV-----LSEAS-EIQCICSVSREBILFSPNLLERAPKIKCGDIH 72  
 Q9548 161 GKLDDLIVVHLNGLSSSSNPYVFNEDPVDGRGKGLVEVLLSLVLAFFPNVFLNRGNH 220  
 Db 73 GSYSDLRLRPFYGGPPEAN YLFLDYVDRGKQSATICLIGYKIKYPNPFLLRGNH 131  
 Q9548 221 EHSMVARYGFREVEVKYPRMKRILAFIDEVYRMPPLASVLSRVLIVGGS-DSTS 279  
 Db 132 EASINRYIYGFDCKRPF--NMRLMKVFTDCFCNLCPVAAIIDKKLICMGGLSDLTH 188  
 Q9548 280 LDLIKSDRGKGVSLRPLPTDGPRLDKTEMQIDIMSPD-QATMGCVNTLRAAGW 338  
 Db 10 INRLLBV-RGRPGKQVQ-----LSEA--EIQCICLVSKDIFMNQPNLKLKAPKIKCG 59  
 Q9548 158 DLKQKDLILVWHLNGLCPSSSPYVFGDFDVRGKGLVEVLLSLVLAFFPNVFLNR 217  
 Db 60 DIHQQYSDLRLRFYGGPPEAN YLFQDYVDRGKQSATICLIGYKIKYPNPFLLR 118  
 Q9548 218 GNHEDSVMARYGFREVEVKYPRMKRILAFIDEVYRMPPLASVLSRVLIVGGS-D 276  
 Db 119 GNHECASINRYIYGFDCKRPF--NMVLKMTDCFCNLCPVAAIIDKKLICMGGLSPE 175  
 Q9548 277 STSLDLIKSDRGKGVSLRPLPTDGPRLDKTEMQIDIMSPD-QATMGCVNTLRAAGW 336  
 Db 176 LHDNQIQNLR-----PCEVPVSGLQDLMNSDPSDSSDVRGMGERSERGS 220  
 Q9548 337 VMECPDVDTNFLQRHLSVIRSHCKPENGHERMDNKITTSASNYTAIGSNKGAYIR 396  
 Db 221 YTGFADRVKEFLQGDLDLICRAHODVGYEFPAQNQVLTFSAPNYCGERFDNAGAMVTD 280  
 Q9548 397 LNNOQMPHVOYISAASOTKRSFKRQMGIVVERSALK 433  
 Db 281 VNEELIVCSR-OQIKELDKKEPKFSGSTTVKSSPTK 316  
 RESULT 27  
 T21322 phosphoprotein phosphatase (EC 3.1.3.16) 1 F25B3.4 [similarity] - *Caenorhabditis elegans*  
 C;Species: *Caenorhabditis elegans*  
 C;Accession: T21322  
 R;Gardner, A.  
 A;Reference number: Z19405  
 A;Accession: T21322  
 A;Status: translated from GB/EMBL/DDJB  
 A;Molecule type: DNA  
 A;Residues: 1-221 <WIL>  
 A;Cross-references: UNIPROT:Q27496; UNIPARC:UPI000007D589; EMBL:Z70752; PIDN:CAA94756.1;  
 A;Experimental source: clone F25B3  
 C;Genetics:  
 A;Gene: C25B3.4  
 A;Map position: 5  
 A;Intonons: 20/2; 15/3; 27/3  
 C;Superfamily: serine/threonine protein phosphatase; phosphoesterase core homology; phosphotyrosine phosphatase homology <PPP>  
 F;7-268/Domain: phosphoprotein phosphatase homology <PPP>  
 F;35-403/Domain: phosphoprotein phosphatase core homology <PPP>  
 F;41-43, 69/Binding site: iron (Asp, His, Asp) #status predicted  
 F;69, 101, 151, 221/Binding site: zinc (Asp, Asn, His, His) #status predicted  
 F;72, 202, 221/Active site: Asp, His, Tyr #status predicted  
 F;73, 200/Binding site: substrate phosphate (Arg) #status predicted





Oy 361 ECKPNQHGFPMHDNKLITIFASNYAIGSNKGAYIRLNQLOMPHPVOYISAASQTKR 417  
 :||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:  
 Db 255 QVVEDGYBFFADQLVTIFSPAPNYCGEFDNAGAMMSVDESNCSP-QILKPADRKPR 310

RESULT 32

T09547

phosphoprotein phosphatase (EC 3.1.3.16) 1, catalytic gamma chain - alfalfa

C;Species: Medicago sativa (alfalfa)

C;Date: 16-Jul-1999 #sequence\_revision 16-Jul-1999 #text\_change 09-Jul-2004

C;Accession: T09547

R;Viselli, E.; Csordas, Toth, B.; Kovacs, I.; Magyar, Z.; Horvath, G.; Bagossi, P.; Gergely Arch. Biochem. Biophys., 360, 206-214, 1998

A;Title: Protein phosphatase 1 catalytic subunit isoforms from alfalfa: Biochemical characterization number: Z16730; MUID:9068922; PMID:9851832

A;Accession: T09547

A;Residues: 1-325 <VIS>

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: mRNA

A;Cross-references: UNIPROT:O65945; UNIPARC:UPI000009DAB2; EMBL:AJ002486; NID:93176073;

A;Experimental source: subspecies sativa, strain RA3

C;Genetics:

A;Gene: PPII gamma

C;Function:

A;Description: catalyzes hydrolysis of peptidyl-phosphoserine or -phosphothreonine to release phosphate; serine/threonine protein phosphatase; phosphoesterase core homology; phosphoprotein phosphatase core homology; phosphotyrosine-specific phosphatase

C;Keywords: phosphoprotein phosphatase; phosphoesterase core homology; phosphotyrosine-specific phosphatase

F;25-84/Domain: phosphoprotein phosphatase homology <PPB>

F;53-121/Domain: Phosphoesterase Core homology <PEC>

Query Match 12.2%; Score 418.5; DB 2; Length 325;  
 Best Local Similarity 30.1%; Pred. No. 5.2e-19; Matches 102; Conservative 66; Mismatches 120; Indels 51; Gaps 9;

Db 7 DIINRLLEV-----RNPRG----KQVQL---SESRTRQLCNWSRDFFLK 43

Oy 90 DIVNAKILPPIRKHNDLIDVKKRKNRHLHVVALLREAKSILQPLNIS----- 143  
 |||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:  
 Db 93 -RWCYVCLARQLQFPMHQMLLEUP-----PLKIGSDHSGPDLURLENAGYPESN 120  
 |||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:  
 Db 52 QMRRKRKNSSSRFGLNIT-----DQLISTLVSVKWPKHTLTDVTEKBI----- 92  
 |||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:  
 Oy 121 HKPVVALILREAKSILQPLNISPVSTAVSQOVTVCDDLGKDDLVWLNKGQLPSSN 180  
 |||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:  
 Db 144 -PVSTAVSQOVTVCDDLGKDDLVWLNKGQLPSSN 180  
 |||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:  
 Oy 44 QPFLLQLEPRKTCGDVQPSQSLRUFYGGFPPEAN-YLFKADYVDRGKQSLETICLI 102  
 |||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:  
 Db 203 LSIYLAEPNAVFLNRGNHDSVMARYGKPTREYESKYNRHLAFIDEVVRWLPGSV 262  
 |||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:  
 Db 103 LAYKIKYPERPENPFLLRGNHCASINRYGPDECKRP-----NVLWKTPTDCSNCLPVAAL 159  
 |||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:  
 Oy 263 LNSRVLIVFGFS-DSTSLDLISDRKSYKVLIRLPPLTGDRLKNEWQFDIMSDP 321  
 ::||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:  
 Db 160 IDEKILCMEGGLSPDHLQDILQRNLR-----PTPDVDTGICLQDWSDP 204  
 |||:||:||:||:||:||:||:||:||:||:||:||:||:||:  
 Oy 322 QATMGCVENTLRLGAGVWFGDPDVTFNLGHRSYVIRSHCKNGHBTMDNKLITFSA 381  
 |||:||:||:||:||:||:||:||:||:||:||:||:||:||:  
 Db 205 SKQVQGWMGRDRCVSYTGKDTIBFLSKHDLICRAHQVVEDGYSFRANQKVITYSA 264  
 |||:||:||:||:||:||:||:||:||:||:||:||:  
 Oy 382 SNTYAIGSNKGAYIRLNQLOMPHPVOYISAASQTKR 420  
 |||:||:||:||:||:||:||:||:  
 Db 265 PNYCGEFDNAGAMMSVDETMCSF-QILKPADKTKLNP 302

RESULT 33

T18936

phosphoprotein phosphatase (EC 3.1.3.16) 1-beta catalytic chain - Caenorhabditis elegans

C;Species: Caenorhabditis elegans

C;Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jul-2004

C;Accession: T18936; T25398

R;Mc Murray, A.; submitted to the EMBL Data Library, June 1996

A;Accession: T18936

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: DNA

A;Cross-references: UNIPROT:P12982; UNIPARC:UPI000012488P; GB:X55198; GB:S47852; NID:98; R;Dombradi, V.; Axton, J.M.; Glover, D.M.; Cohen, P.T.W.

A;Residues: 1-401 <WIL>

RESULT 34

PYPAFA

phosphoprotein phosphatase (EC 3.1.3.16) 1-alpha-2 catalytic chain - fruit fly (Drosophila)

C;Species: Drosophila melanogaster

C;Date: 31-Mar-1992 #sequence\_revision 31-Mar-1992 #text\_change 05-Oct-2004

C;Accession: S12960; S05524

R;Dombradi, V.; Axton, J.M.; Barker, H.M.; Cohen, P.T.W.

FEBS Lett. 275, 39-43, 1990

A;Title: Protein phosphatase 1 activity in Drosophila mutants with abnormalities in mitc

A;Reference number: S12960

A;Accession: S12960; MUID:91085574; PMID:2175717

A;Residues: 1-302 <DN>

A;Cross-references: UNIPROT:P12982; UNIPARC:UPI000012488P; GB:X55198; GB:S47852; NID:98; R;Dombradi, V.; Axton, J.M.; Glover, D.M.; Cohen, P.T.W.

RESULT 35

Query Match 12.0%; Score 412; DB 2; Length 316;  
 Best Local Similarity 30.6%; Pred. No. 1.3e-18;  
 Matches 100; Conservative 61; Mismatches 112; Index 54; Gaps 10;  
 Matches 100; Conservative 61; Mismatches 112; Index 54; Gaps 10;

OY	90 DIVNAKTELPTRKHNIDLLIDVFRKK--GMRLLHPKVALLTRRAKSLKLQLPNTS---	143
Db	24 DIRRLIVE-----FRNTRPGSGKQHV-----LSEG--EIROLCAWSKD1	60
OY	144 ---PVSTAVSQVTVGSDIHKQLDILWVHKNGLPPSSNPYVPGNDFVPGKGRERFLV 199	
Db	61 FLQQPILLEAPLAKICDINGQSYSLRFLFFYGGFPDAN-YLFGLDYVPGKQLETT1 119	
OY	200 LLLSLVIAFPNVAFLNRGNHEDSVMMARYGRREVESKYPRNHKRLAFIDEVRMLPL 259	
Db	120 CFIILAKVKKYPENPFLANGNHCASINRIGYGRERFLR--NVRLWKFIDCPNLPLV 176	
OY	260 GSVLNSRVTLIVHGGS-DSTSLSLKSIDRKGVKVSILRPLPLDGPBLDKTQSQIDMM 318	
Db	177 AALID-RILCMEHGEGISFPLMSLQDQNSR-----PLDIPDSGLVCDLW 220	
OY	319 SDPQATMOCVPTVLRGAGGWWPPDVTNFLPHRLISVVRSHCKPKHGRHEFHNDKLT1 378	
Db	221 SDPSGDYVKGWGANDGRGSVYTFGRADTVFLQFLQNDMLICRAHQVQVSDQYEFPAERQVTV 280	
OY	379 FGSNTYIAIGSNKGAYIRLNNOIAMPHF 405	
Db	281 FSAPNYCGEPFDNAGAMMSIDESLWMCSP 307	

RESULT 36

PABY3  
 phosphoprotein phosphatase (EC 3.1.3.16) PPH3 - yeast (*Saccharomyces cerevisiae*)  
 N:Alternative names: protein D4421; protein YDB554.08; protein YDR075W  
 C:Species: *Saccharomyces cerevisiae*  
 C:Date: 30-Jun-1993 #Sequence revision 10-Feb-1995 #text\_change 05-Oct-2004  
 C:Accession: S44331; S47672; S55822; S49830; C41555; S67892; S17483  
 R:Hoffmann, R.; Jung, S.; Bhrmann, M.; Hofer, H.W.  
 Yeast 10, 567-578, 1994  
 A:Title: The *Saccharomyces cerevisiae* gene PPH3 encodes a protein phosphatase with prop  
 A:Reference number: S44331; PMID:95028155; PMID:7941742  
 A:Accession: S44331  
 A>Status: not compared with conceptual translation  
 A:Molecule type: DNA  
 A:Residues: 1-308 <HOB>  
 A:Cross-references: UNIPROT:PI32345; UNIPARC:UPI0000053121  
 R:Coster, F., Jonniaux, J.L., Goffeau, A.  
 submitted to the EMBL Data Library, October 1994  
 A:Reference number: S48758  
 A:Accession: S48762  
 A:Molecule type: DNA  
 A:Residues: 1-308 <COS>  
 A:Cross-references: UNIPARC:UPI0000053121; EMBL:X82086; NID:9558241; PIDN:CAA57602.1; P  
 R:Coster, F., Jonniaux, J.L., Goffeau, A.  
 Yeast 11, 673-679, 1995  
 A:Title: Analysis of a 32.8 kb segment of yeast chromosome IV reveals 21 open reading f  
 A:Accession number: S55819; MURID:96093910; PMID:7483840  
 A:Status: nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-308 <COS>  
 A:Cross-references: UNIPARC:UPI0000053121; EMBL:X82086; NID:9558241; PIDN:CAA57602.1; P  
 A:Note: the nucleotide sequence was submitted to the EMBL Data Library, October 1994  
 R:Richards, C., Harris, D.B.  
 submitted to the EMBL Data Library, November 1994  
 A:Reference number: S49823  
 A:Accession: S49830  
 A:Molecule type: DNA  
 A:Residues: 1-308 <RIC>  
 A:Cross-references: UNIPARC:UPI0000053121; EMBL:Z46796; NID:9577794; PIDN:CAA86797.1; P  
 R:Ronne, H.; Carlberg, M.; Hu, G.Z.; Neulin, J.O.  
 Mol. Cell. Biol. 11, 4876-4884, 1991  
 A:Title: Protein phosphatase 2A in *Saccharomyces cerevisiae*: effects on cell growth and  
 submitted to the EMBL Data Library, November 1994  
 A:Reference number: S49823  
 A:Accession: S49830

A; Accession: C41525  
A; Molecule type: DNA  
A; Residues: 1-244, 'I', 246-308 <ROM>  
A; Cross-references: UNIPARC:UPI000016099C; GB:X58858; NID:94211; PIDN:CRA41662.1; PID:94211  
R; Pouri, F.; Jonniaux, J.-L.; Purnelle, B.; Cosper, P.; Goffeau, A.  
Submitted to the Protein Sequence Database, July 1996  
A; Reference number: S67892  
A; Accession: S67892  
A; Molecule type: DNA  
A; Residues: 1-308 <FOU>  
A; Experimental source: strain S288C  
A; Genetics:  
A; Gene: SGD:PPH3; MIPS:YDR075W  
A; Cross-references: SGD:S0002482; MIPS:YDR075W  
A; Map position: 4R  
C; Function:  
A; Description: phosphoric monoester hydrolase  
C; Superfamily: serine/threonine protein phosphatase; phosphoesterase core homology; phosphatase homology; phosphotriesterase homology  
C; Keywords: iron; metalloprotein; phosphoric monoester hydrolase; serine/threonine-specific  
F; 17-17/Domain: phosphoprotein phosphatase homology <PP>  
F; 45-113/Domain: phosphocesterase core homology <PEC>  
F; 51 53 79/Binding site: iron (Asp, Asn, His, Asp) #status predicted  
F; 82, 111, 161, 235/Binding site: zinc (Asp, Asn, His, His) #status predicted  
F; 82, 112, 259/Active site: Asp, His, Tyr #status predicted  
P; 83, 208/Binding site: substrate phosphate (Arg) #status predicted

Query Match Score 12.0%; Score 411.5; DB 1; Length 333;  
Best Local Similarity 31.0%; Pred. No. 1.5e-18; Mismatches 65; Mismatches 133; Indels 25; Gaps 8;  
Matches 100; Conservative 56; Mismatches 111; Indels 25; Gaps 8;

Qy 128 ILREAKSKIQQLPISPVSTAVSQVPTVCGDPLKGLDPLWVHLNGKLPSSSNPVYENPDVDEGKGRLGEVLLILSLXLAFFAVP 214  
Db 22 VFRILCLNSQBLMMEGNV-TQDTPPTVTCGDDQQLHDLITLFKSGCVERTR-YIRLD 79  
Qy 188 FVDRGKRGKIEFLVLLSLXLAFFAVHLLNGHDSVWARKVIREKESKPKRNKHL 247  
Db 80 FVDRGKPSHSRSPLLCYKRYPDIRTIRGNHETRQITKVKYGPYDVKY--GSNUW 137  
Qy 248 AFIDEVYWRMLPLGSVTLNSRVLIVGGFS-DTSLSLDSIKSDRGKVSVTLPPLGEPLD 306  
Db 138 RYCCEVFPYLSIGLIIINNSICVHGGLSDFMTWDBITDRKQBP---HEG--- 187  
Qy 307 KTEWQKQIFPIMMWSPQ-ATMGCVPTNLARGAGWFGPPVTDNFQERHLSYVIRSHECKP 364  
Db 188 ----AMCDLWLSPEDVUTWLSLSP--RGAGFLFGKREBDQFLKNNVLLARAHOLW 239  
Qy 365 NGHFPMHDKILITFSASGTYAOSNKGYVRLQMLMPHWQVISAQSQK 416  
Db 240 EGYKEMPGDGLVTWSAPVPCYRCGVNAVAVLKDDLNRRYTFEAQVQNE 291

RESULT 37

T16476 phosphoprotein phosphatase (EC 3.1.3.16) 1-alpha F56C9.1 [similarity] - Caenorhabditis elegans

N; Alternative names: protein phosphatase 1, catalytic subunit

C; Species: *Caenorhabditis elegans*

C; Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 05-Oct-2004

C; Accession: T16476

R; Du, Z  
Submitted to the EMBL Data Library, May 1994

A; Description: the sequence of C. elegans cosmid F56C9.

A; Reference number: S4729

A; Accession: T16476

A; Status: translated from GB/EMBL/DDJB

A; Molecule type: DNA

A; Residues: 1-382 <WIL>

A; Cross-references: UNIPARC:UPI0000172815; EMBL:Z46996; PIDN:CAA87100.1; GSPDB:GN00021; C; Genetics:

A; Gene: CESP:C34C12.3

A; Map position: 3

A; Intros: 77/2; 107/2; 145/1; 179/2; 221/3; 265/3; 292/3; 336/1

C; Superfamily: serine/threonine protein phosphatase; phosphoesterase core homology; phosphatase homology; phosphotriesterase homology

C; Keywords: iron; metalloprotein; phosphoric monoester hydrolase; serine/threonine-specific

F; 96-356/Domain: phosphoprotein phosphatase core homology <PP>

F; 124-192/Domain: phosphocesterase core homology <PEC>

F; 130, 132, 158/Binding site: iron (Asp, Asn, His, Asp) #status predicted

F; 158, 190, 240, 315/Binding site: zinc (Asp, Asn, His, His) #status predicted

F; 161, 191, 339/Active site: Asp, His, Tyr #status predicted

P; 162, 288/Binding site: substrate phosphate (Arg) #status predicted

Query Match Score 11.9%; Score 409.5; DB 1; Length 382;  
Best Local Similarity 37.3%; Pred. No. 2.4e-18; Mismatches 98; Indels 25; Gaps 7;  
Matches 100; Conservative 45; Mismatches 98; Indels 25; Gaps 7;

Qy 141 NISPVSTAVSQVPTVCGDPLKGLDPLWVHLNGKLPSSSNPVYENPDVDRKGRLGEVLL 200  
Db 117 NWVPSV---PTVCCDINGPFDI-BLKFGGTWPNKTVFMDYVDRKHSRHTVT 171

C; Keywords: glycogen metabolism; heterodimer; iron; metalloprotein; phosphoric monoester

F; 29-208/Domain: phosphoprotein phosphatase homology <PP>

F; 57-155/Domain: phosphoesterase core homology <PEC>

F; 63, 65, 91/Binding site: iron (Asp, Asn, His, Asp) #status predicted

F; 91, 123, 172, 247/Binding site: zinc (Asp, Asn, His, His) #status predicted

F; 94, 124, 271/Active site: Asp, His, Tyr #status predicted

Query Match Score 12.0%; Score 411.5; DB 1; Length 333;  
Best Local Similarity 31.0%; Pred. No. 1.5e-18; Mismatches 65; Mismatches 133; Indels 25; Gaps 8;  
Matches 100; Conservative 56; Mismatches 111; Indels 25; Gaps 8;

Qy 98 LPTRKNHIDLIDLVFRKGERGNRLRHKVVALI--LRRAKSILKOLPNISPVSTAVSPOOT 154  
Db 1 MDVEKLNDNNIISRLLSVGRSK-PGKVNOLTESBESIKLGCKSRIBISQPLIBLEAPLK 59  
Qy 155 VCDPLHAKDPLWVHLNGLSSSNPVYENPDVDEGKGRLGEVLLILSLXLAFFAVP 214  
Db 60 ICIDVH3OYYDVLQDPLRVEGPGPPBESN-YLFGKQDPLRQKQSLTICLAVKLYPKRNP 118  
Qy 215 LNRGENHEDSUMMARYTREVESKYPHNKRILAPIDEVYRMLPLGSVLSRVLIVGGP 274  
Db 119 LURGNHESASINRIVYGFDECERY-NNIKWKFTP-CFNCLPQVAlAIIIDRKICFCIGGL 175  
Db 275 S-DSTSALIJKSIDSQKVVSLRPLPLGEPLDKTBTWQDPIQDMWSPQATMCVPHLTLR 333  
Db 176 SPDLQSMQEQRIRRMR-----PTVDPDGICLQDILWSPDKDVTGWGENDR 220  
Qy 334 GACWVWRPAPDVFLQHRLSVIRSHICKPKHGRPHRDNKKTIFASNYVAIGSKGA 393  
Db 221 GVSFTPGRPEPVAKPLHKKDLDLICRAHQVVEDGYEPFAKQLOVLTFSAPNVCGBEDNAGS 280  
Qy 394 YIRLNNOIMPHFQYQIAASQTK 416  
Db 281 MWVTDLMSCF-QILKPADK 302

RESULT 38

T19701 phosphoprotein phosphatase (EC 3.1.3.16) C34C12.3 - Caenorhabditis elegans

C; Species: *Caenorhabditis elegans*

C; Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 05-Oct-2004

R; Kershaw, J.  
Submitted to the EMBL Data Library, December 1994

A; Reference number: Z19166

A; Accession: T19701

A; Status: translated from GB/EMBL/DDJB

A; Molecule type: DNA

A; Residues: 1-382 <WIL>

A; Cross-references: UNIPARC:UPI0000172815; EMBL:Z46996; PIDN:CAA87100.1; GSPDB:GN00021; C; Genetics:

A; Gene: CESP:C34C12.3

A; Map position: 3

A; Intros: 77/2; 107/2; 145/1; 179/2; 221/3; 265/3; 292/3; 336/1

C; Superfamily: serine/threonine protein phosphatase; phosphoesterase core homology; phosphatase homology; phosphotriesterase homology

C; Keywords: iron; metalloprotein; phosphoric monoester hydrolase; serine/threonine-specific

F; 96-356/Domain: phosphoprotein phosphatase core homology <PP>

F; 124-192/Domain: phosphocesterase core homology <PEC>

F; 130, 132, 158/Binding site: iron (Asp, Asn, His, Asp) #status predicted

F; 158, 190, 240, 315/Binding site: zinc (Asp, Asn, His, His) #status predicted

F; 161, 191, 339/Active site: Asp, His, Tyr #status predicted

P; 162, 288/Binding site: substrate phosphate (Arg) #status predicted

Query Match Score 11.9%; Score 409.5; DB 1; Length 382;  
Best Local Similarity 37.3%; Pred. No. 2.4e-18; Mismatches 98; Indels 25; Gaps 7;  
Matches 100; Conservative 45; Mismatches 98; Indels 25; Gaps 7;

Qy 141 NISPVSTAVSQVPTVCGDPLKGLDPLWVHLNGKLPSSSNPVYENPDVDRKGRLGEVLL 200  
Db 117 NWVPSV---PTVCCDINGPFDI-BLKFGGTWPNKTVFMDYVDRKHSRHTVT 171

RESULT 39

QY 201 LLLSLYLAPPNAVFLNRGNHEDSVNARYFIREVESKYPPIRKRLAFIGDLYRMLPLG 260  
 Db 172 LRCQLLKVPPNQTLTGRHESRSRISNYGFDCECQNY--GHGVHVKPCKVFDVLPIG 229  
 QY 261 SVLNRSVLVTHGGRS-DSTSLDIKSDIKRKGKVSILAPPLTGEPBLDKTEWQIFDIMS 319  
 Db 230 ALIDESVLVHGGSLSPDRTRIDSMILURAOQEVPNKGP---LCDIMWS 274  
 QY 320 DPOATMGCYPTNLRGACTWFGPQUTDNFLQRHLSYVRHSCKPNGHFRPMEDNKITIP 379  
 Db 275 DPDDDVEDWVISORGAGKVFPGAKVTEEFIMMNDLILCRSHOLVDGPKYMFNEKLATW 334  
 QY 380 SASNY-YAIGNSKGAY-IRLNNOLMPHF 405  
 Db 335 SAPNYCYRCGNAAAVFEIDGNNRSTKPF 362

C;Species: Arabidopsis thaliana (mouse-ear cress)  
 C;Date: 13-Jan-1995 #sequence\_revision 13-Jan-1995 #text\_change 09-Jul-2004  
 C;Accession: S31089; T12966  
 R;Smith, R.D.; Walker, J.C.  
 Plant Mol. Biol. 21, 307-316, 1993  
 A;Title: Expression of multiple type I phosphoprotein phosphatases in Arabidopsis thaliana  
 A;Reference number: S31089; MUID:93144705; PMID:7678768  
 A;Molecule type: mRNA  
 A;Residues: 1-312 <SMI>  
 A;Cross-references: UNIPROT:P48485; UNIPARC:UPI0000131FB5; EMBL:M93412; NID:916602; PIR:R.Cholne, N.; Robert, C.; Broquier, P.; Winkler, F.; Cattolico, L.; Antignan, P.; Se  
 Submitted to the Protein Sequence Database, July 1999  
 A;Accession: T12966  
 A;Reference number: 217586  
 A;Molecule type: DNA  
 A;Residues: 1-312 <CHO>  
 A;Cross-references: UNIPARC:UPI0000131FB5; EMBL:AL096859; GSFDB:GN00061; ATSP:T6H20, 150  
 A;Experimental source: cultivar Columbia; BAC clone T6H20  
 C;Genetics:  
 A;Gene: ATSP:T6H20..150  
 A;Map position: 3  
 A;Introns: 69/1, 255/3  
 C;Superfamily: Serine/threonine protein phosphatase; phosphoesterase core homology; phosphatase  
 C;Keywords: iron; metalloprotein; phosphoric monoester hydrolase; serine/threonine-specific  
 P;27-286//Domain: phosphoprotein phosphatase homology <PPC>  
 P;55-123//Domain: phosphoesterase core homology <PEC>  
 A;Molecule type: mRNA  
 A;Residues: 1-316 <SMI>  
 A;Cross-references: UNIPROT:P22199; UNIPARC:UPI0000131FB5; EMBL:M60215; NID:9168722; PID:91226; PIR:R.Cholne, N.; Robert, C.; Broquier, P.; Winkler, F.; Cattolico, L.; Antignan, P.; Se  
 Submitted to the Protein Sequence Database, July 1999  
 C;Superfamily: Serine/threonine protein phosphatase; phosphoesterase core homology; phosphatase  
 C;Keywords: iron; metalloprotein; phosphoric monoester hydrolase; serine/threonine-specific  
 P;70-72, 98//Domain: phosphoprotein phosphatase homology <PPC>  
 P;61-63, 89//Domain: phosphoesterase core homology <PEC>  
 P;89, 121-170, 245//Binding site: zinc (Asp, Asn, His, His) #status predicted  
 P;92, 122, 269//Active site: Asp, His, Tyr #status predicted  
 P;93, 218//Binding site: substrate phosphate (Arg) #status predicted

Query Match Similarity 11.9%; Score 409; DB 2; Length 316;  
 Best Local Similarity 30.1%; Pred. No. 2.3e-18;  
 Matches 101; Conservative 66; Mismatches 127; Indels 42; Gaps 10;

Query Match Similarity 11.9%; Score 408; DB 2; Length 312;  
 Best Local Similarity 30.1%; Pred. No. 2.3e-18;  
 Matches 107; Conservative 63; Mismatches 118; Indels 58; Gaps 10;

Query Match Similarity 11.9%; Score 408; DB 2; Length 312;  
 Best Local Similarity 30.1%; Pred. No. 2.3e-18;  
 Matches 107; Conservative 63; Mismatches 118; Indels 58; Gaps 10;

Db 4 OQQGSMHDPAVADD-----IIRR----LIDYRNPKAGIK----QA 34

QY 128 IIREAAMSILKOLNPNTS-----PVSTAVSQVQTYCGDLHGKLDLWVWIKNGLPSNN 180  
 Db 35 MLDNS--EIRQCFVSRBIFLQOQPCLBLAPVKICGDINGHQYSBILRLFFYGGFPAA 92

QY 181 PYVPGEDFVDGKGKGLVELLILSLYLAPPNAVFLNAR 217  
 Db 61 DVHQGYSQSLRFLDFYGGPPQAN-YLFQGDYVDRGKOSLETICLLAYKVKYPENPFLR 119

QY 218 GHNFDSVWANAFIREVEKSPRNNKKLAFIDEVTRWLPGSVLNRSVLVHGRS-D 276  
 Db 93 -YLFQGDYVDRGKOSLETICLLAYKVKYPENPFLRGNHECASTINRYKSFYDECKRRF- 150

QY 241 RNHKRRLAFIGDVEVYKMLPLGSVLSVNLVHGGFS-DSTSLDIKSDIKRKGKVSILRPL 299  
 Db 151 -NVKLMKVFDT-TENCPVAAVTDKLCMGLSPBLINVRQNIER----- 197

QY 300 TDGEPLKTEWQIQFIDMWSDPQATMGCVPTLRGAGGWVFGDVTDNFLORHLSVIRS 359  
 Db 198 ---PTVVPDGICLILMSDPSKDVKGWGMNDRGVSYTFGADKVAEFLIKNDMDLVCR 253

QY 360 HRCKPKNGHEFMHDNKITITFASNSNTAIGNSKGAY 405  
 Db 220 VSPTFGPDVKVSEPLAKDILICRAHOVQEFGYEFFASRQVLTFSAPNYCQRFNDNAGAM 279

QY 395 IRANQOLMPHF-----VOYISAQSQTKLISPKORMG 425  
 Db 280 MSVDTTLMCSCFQILPKARKMMGSGSTNNKSGFKSFRG 315

RESULT 40

S31089 Phosphoprotein phosphatase (EC 3.1.3.16) 1 catalytic chain (clone TOP5) - Arabidopsis t  
 N;Alternate names: protein T6H20..150

Search completed: January 20, 2006, 19:52:41  
 Job time : 42 secs

Gencore version 5.1.6  
 Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - protein search, using SW model

Run on: January 20, 2006, 19:38:07 ; Search time 163 seconds

Sequence: 1 MDENATRAAIFIQKWMRRHQ..... . . . . . VEHDIDPTDCESKVDPKKS 661

Title: US-09-463-733-1  
 Perfect score: 3442  
 Scoring table: BLOSUM62  
 Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : UniProt\_05\_80;\*

- 1: uniprot\_sprot:\*
- 2: uniprot\_trembl:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Match Length	DB ID	Description
1	3442	100.0	661	1	RDG DROME
2	3121	90.7	603	2	05UD3_DROME
3	1930.5	56.1	613	2	07Q29_ANOGCA
4	1813	52.7	604	2	07Q01_ANOGCA
5	1371.5	39.8	705	2	04S0B1_TETRIG
6	1359	39.5	757	1	PPE2_MOUSE
7	1331	39.0	707	2	061650_CABER
8	1330	38.6	753	1	PPE2_HUMAN
9	1317	38.3	707	2	001921_CASELL
10	1285.5	37.3	653	2	04R4Y0_MACPA
11	1271.5	36.9	653	1	PBP1_HUMAN
12	1263	36.7	683	2	09W674_FUGU
13	1238.5	36.0	671	2	04SHR7_TETRIG
14	1232	35.8	797	2	04R011_TETRIG
15	1147	33.3	572	2	08MW12_CASELL
16	839	24.4	405	2	04R441_MACPA
17	627.5	18.2	493	2	068BPO_XENR
18	618.5	18.0	897	2	04Y018_PLACH
19	615.5	17.9	492	2	042205_XENMA
20	615.5	17.9	493	2	06GP56_XENDA
21	614.5	17.9	493	2	07QB11_ANOGCA
22	613	17.8	874	2	04RFB2_TETRIG
23	603.5	17.5	953	2	04Y018_PLACH
24	594.5	17.3	923	2	081728_TRICR
25	582.5	16.9	473	2	0512PS_ENTH
26	581	16.9	496	2	09NESB_CASEL
27	579	16.8	829	2	04Z5E3_PLABR
28	577.5	16.8	594	2	08W0R3_PLAPKA
29	577.5	16.8	658	2	081DE7_PLAP7
30	577	16.8	522	2	09VB11_DRONE
31	575.5	16.7	525	2	05CJA8_CRYPTO

RESULT 1	RDG DROME	STANDARD;	PRT;	661 AA.
ID	RDG DROME			
AC	P40421; Q9WMA4;			
DT	01-FEB-1995 (Rel. 31, Created)			
DT	01-FEB-1995 (Rel. 31, Last sequence update)			
DT	13-SRP-2005 (Rel. 48, Last annotation update)			
DE	"Drosophila retinal degeneration C (rdgC) encodes a novel serine/threonine protein phosphatase.";			
GN	Name=rdgC; ORFNames=CG6571';			
OS	Drosophila melanogaster (Fruit fly).			
RA	Bukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Sphyradoidea; Drosophilidae; Drosophila.			
OC				
OX	NCBI_TaxID:7227;			
RN				
RX	NUCLEOTIDE SEQUENCE (ISOFORM B); MBDLNRB-92266398; PubMed=1316807; DOI=10.1016/0092-8674(92)90230-A; Steel, P.R., Washburn, T., Rieger, R., O'Tousa, J.B.; "Drosophila retinal degeneration C (rdgC) encodes a novel serine/threonine protein phosphatase.";			
RU	Cell 69:669-676(1992).			
[2]				
RP	NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].			
RC	STRAIN:Berkeley;			
RX	MEDLINE=2016006; PubMed=10731132; DOI=10.1126/science.287.5461.2185; Adams, M.D., Celinski, S.E., Holt, R.A., Evans, C.A., Gocayne, J.D., Amanatides, P.G., Scheer, S.E., Li, P.W., Hoskins, R.A., Gille, R.P., George, R.A., Lewin, B.S.E., Richardson, S., Ashburner, M., Henderson, S.N., Sutton, G.G., Lewellen, J.R., Yandell, M.D., Zhang, Q., Chen, L.X., Brandon, R.C., Rogers, Y.-H.C., Blazquez, R.G., Champagne, M., Pfeiffer, B.D., Wan, K.H., Doyle, C.J., Baxter, G.E., Heit, G., Nelson, C.R., Miklos, G.L.G., Abrial, J.-F., Agbavani, A., An, H.-J., Andrews-Pfannkoch, C., Baldwin, D., Ballew, R.M., Basu, A., Baxendale, J., Bayraktaroglu, L., Beasley, E.M., Bebbon, K.Y., Benos, P.V., Berman, B.P., Bhandari, D., Bolshakov, S., Borckova, D., Botchkina, M.R., Bouck, J., Brokstein, P., Brottner, P., Burtis, K.C., Busam, D.A., Butler, H., Cadieu, E., Center, A., Chandra, I., Cherry, J.M., Cowley, S., Dahlke, C., Davenport, L.B., Davies, P., de Pablo, B., Deicher, P., Deng, Z., Mays, A.D., Dew, I., Dietz, S.M., Dodson, K., Doupe, L.E., Downes, M., Dugan-Rocha, S., Dunkov, B.C., Dunn, P., Dubrin, K.J., Evangelista, C.C., Ferraz, C., Ferreira, S., Fleischmann, W., Fosler, C., Gabril, A.B., Garg, N.S., Geibert, W.M., Glaser, K., Glodek, A., Gong, F., Gorrell, J.H., Gu, Z., Guan, P., Harris, M., Harris, N.L., Harvey, D.A., Heiman, T.J., Hernandez, J.R., Houck, J., Houston, K.A., Howland, T.J., Wei, M.H., Ibegwam, C., Jalali, M., Kalush, F., Karpen, G.H., Ke, Z., Kemison, J.A., Ketchum, K.A., Kimmel, B.E., Kodilla, C.D., Kraft, C., Kravitz, S., Kuld, D., Lai, D., Li, Z., Liang, Y., Lin, X., Liu, X., Matrei, B., McIntosh, T.C., McLeod, M.P., McPherson, D., Merkulov, G., Milshina, N.V., Moharry, C., Morris, J., Moshrefi, A., Mount, S.M., Moy, M., Murphy, B., Muzzey, D.M., Nelson, D.L., Nelson, D.R., Nelson, K.A., Nixon, K., Nusskern, D.R., Paclob, J.M., Palazzo, M., Pittman, G.S., Pan, S., Pollard, J., Purif, V., Reece, M.G., Reinert, K., Remington, K., Saunders, R.D.C., Schueler, F., Shen, H.,			

Oufsf6 theileria a  
 Q662n plasmidium  
 Q9w6j3 oryza sativ  
 Q9lh4 lycopersico  
 Q9pw0 homo sapien  
 PR3041 homo sapien  
 Q93fr0 homo sapien  
 Q93xv homo sapien  
 Q94538 rattus norvegicus  
 Q90676 rattus norvegicus  
 PR3042 rattus norvegicus  
 Q9616 rattus norvegicus  
 Q958t2 ponga pygmaea  
 Q9pix6 homo sapien

RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
 RA Spier B., Spradling A.C., Stapleton M., Strong R., Sun S.,  
 RA Svoboda R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
 RA Wang Z.-Y., Wasserman D.A., Weinstock G.M., Weissbach J.,  
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
 RA Yeh J.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
 RT "The genome sequence of *Drosophila melanogaster*."  
 RL Science 287:2185-2195(2000).  
 RN [3]  
 RP GENOME REANNOTATION AND ALTERNATIVE SPlicing.  
 RX MEDLINE=22426069; PubMed=12537572;  
 RA Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,  
 RA Bradecy P., Huang Y., Kanamori J.S., Milburn G.H., Prochnik S.E.,  
 RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Bernan B.P.,  
 RA Bettencourt B.R., Celinski S.B., de Grey A.D.N.J., Drysdale R.A.,  
 RA Harris N.L., Richter J., Risso S., Schrodter A.J., Shu S.Q.,  
 RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,  
 RT Lewis S.E.;  
 RT "Annotation of the *Drosophila melanogaster* euchromatic genome: a  
 systematic review."  
 RL Genome Biol. 3:RESEARCH0083.1-RESEARCH0083.22(2002).  
 CC -I- FUNCTION: Phosphatase required to prevent light-induced retinal  
 degeneration.  
 CC -I- CATALYTIC ACTIVITY: A phosphoprotein + H<sub>2</sub>O = a protein +  
 phosphate.  
 CC -I- COFACTOR: Binds 1 iron ion per subunit (By similarity).  
 CC -I- COFACTOR: Binds 1 manganese ion per subunit (By similarity).  
 CC -I- ENZYME REGULATION: Regulated by calcium.  
 CC -I- ALTERNATIVE PRODUCTS:  
 Event=Alternative splicing; Named isoforms=2;  
 Name=B; Synonyms=C;  
 IsoId=P40421-1; Sequence=Displayed;  
 Name=A;  
 IsoId=P40421-2; Sequence=ISP 009324;  
 Note=No experimental confirmation available;  
 CC -I- TISSUE SPECIFICITY: Expressed in the visual systems of the fly, as  
 well as in the mushroom bodies of the central brain.  
 CC -I- SIMILARITY: Belongs to the PPP phosphatase family.  
 CC -I- SIMILARITY: Contains 3 EF-hand domains.  
 CC -I- SIMILARITY: Contains 1 IQ domain.  
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation at  
 the European Bioinformatics Institute. There are no restrictions on its  
 use as long as its content is in no way modified and this statement is not  
 removed.  
 CC  
 EMBL: M89628; AAC001734.1; -; Genomic DNA.  
 DR EMBL: AE00514; AAFA4044.2; -; Genomic DNA.  
 DR EMBL: AE00514; AA041217.1; -; Genomic DNA.  
 PIR: A42287; A42287.  
 HSSP: Q08209; IAU1.  
 DR Ensembl; CG5571; *Drosophila melanogaster*.  
 DR FlyBase; FBgn0004366; rdgc.  
 DR GO: GO:000516; F:calmodulin binding; IDA.  
 DR GO: GO:001972; P:calcium-mediated signaling; TAS.  
 DR GO: GO:0016059; P:deactivation of rhodopsin mediated signaling; IMP.  
 DR GO: GO:007602; P:phototransduction; IMP.  
 GO; GO:0006470; P:protein amino acid dephosphorylation; IMP.  
 InterPro; IPR01192; EF-hand type.  
 InterPro; IPR00248; EF-hand Ca<sup>2+</sup>-bd.  
 Int-Pro; IPR00048; IQ\_Cam\_Ed\_Region.  
 InterPro; IPR00443; M-peptidase.  
 InterPro; IPR01208; PPBP.  
 InterPro; IPR005186; T\_phthase\_apAH.  
 Pfam; PF00056; eihand; 3.  
 Pfam; PF00612; IQ; 1.  
 Pfam; PF0149; Metallophos; 1.  
 PRINTS; PR00114; STPHPTASE.  
 ProDom; PD000012; EF-hand; 1.

DR PRODom; PD000252; T\_phthase\_apAH; 1.  
 DR PROSTE; PS00018; EF\_HAND\_I; 2.  
 DR PROSTE; PS50222; EF\_HAND\_2; 3.  
 DR PROSTE; PS50066; IQ; 1.  
 DR PROSTE; PS00125; SER\_THR\_PHOSPHATASE; 1.  
 KW Alternative splicing; Calcium; Hydrolase; Iron; Magnesium; Manganese; Metal-binding; Protein phosphatase; Repeat; Sensory transduction; Vision.  
 KW DOMAIN 7 32 IQ.  
 FT DOMAIN 441 476 EF-hand 1.  
 FT DOMAIN 526 561 EF-hand 2.  
 FT DOMAIN 566 601 EF-hand 3.  
 FT CA\_BIND 539 550 1 (potential).  
 FT CA\_BIND 579 590 2 (potential).  
 FT REGION 105 413 Catalytic.  
 FT ACT\_SITE 220 220 Proton donor (By similarity).  
 FT METAL 158 158 Iron (By similarity).  
 FT METAL 160 160 Iron (By similarity).  
 FT METAL 187 187 Iron and manganese (By similarity).  
 FT METAL 219 219 Manganese (By similarity).  
 FT METAL 271 271 Manganese (By similarity).  
 FT METAL 360 360 Manganese (By similarity).  
 FT VARSPIC 1 58 Missing (in isoform A).  
 SQ SEQUENCE 661 AA; 7551 MW; A3DC42933B4CCR3 CRC64;  
 Query Match 100.0%; Score 3442; DB 1; Length 661;  
 Best Local Similarity 100.0%; Prev. No. 2e-2000;  
 Matches 661; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MDENATRAPIQKWRQRHARREMRQRCWQIFONLYEASQDQALYKEFNDLJIKMP 60  
 Db 1 MDENATRAPIQKWRQRHARREMRQRCWQIFONLYEASQDQALYKEFNDLJIKMP 60  
 QY 61 QAGGRNQYOSAHYVSLDDKDLYEFQGIVNAKLEPLTRKNHDLLIDLPVKRGNR 120  
 Db 61 QAGGRNQYOSAHYVSLDDKDLYEFQGIVNAKLEPLTRKNHDLLIDLPVKRGNR 120  
 QY 121 HKPVVALLIREAKSLQKLEMISPVSTAVSQVTYQGHDQKLDDLVVHKGQLPSSN 180  
 Db 121 HKPVVALLIREAKSLQKLEMISPVSTAVSQVTYQGHDQKLDDLVVHKGQLPSSN 180  
 QY 181 PYVFNEDDFVDRGKRQLEVLLSIALAFPAVPLNGTNEQDSVANAYQFIRETESK 240  
 Db 181 PYVFNEDDFVDRGKRQLEVLLSIALAFPAVPLNGTNEQDSVANAYQFIRETESK 240  
 QY 241 RNNKRITIAPIDEVYRMLPLGSLSVNSRVLVLTIGGSDSTSLLDKLISDRGKVSLIRPLT 300  
 Db 241 RNNKRITIAPIDEVYRMLPLGSLSVNSRVLVLTIGGSDSTSLLDKLISDRGKVSLIRPLT 300  
 QY 301 DGPRLDTENKQIFDMMWSPQATMCPVTLRGGWCPDPDVTFQHRLSVTRSH 360  
 Db 301 DGPRLDTENKQIFDMMWSPQATMCPVTLRGGWCPDPDVTFQHRLSVTRSH 360  
 QY 361 ECKPNHSEPMADNKLTITPSASNTYAIAGSNKGAYRLLNQMLPHTVQYSAASQTKRLSF 420  
 Db 361 ECKPNHSEPMADNKLTITPSASNTYAIAGSNKGAYRLLNQMLPHTVQYSAASQTKRLSF 420  
 QY 421 KORMGTVESSAKELAVERMRDRHDLSELDEFPKYDPDKPSGYTISHWKUNENVTGLP 480  
 Db 421 KORMGTVESSAKELAVERMRDRHDLSELDEFPKYDPDKPSGYTISHWKUNENVTGLP 480  
 QY 481 ECKPNHSEPMADNKLTITPSASNTYAIAGSNKGAYRLLNQMLPHTVQYSAASQTKRLSF 420  
 Db 481 ECKPNHSEPMADNKLTITPSASNTYAIAGSNKGAYRLLNQMLPHTVQYSAASQTKRLSF 420  
 QY 481 RLLRKDKLAPGTSOKYNNRTIDLDLTDVTILEKADGMSTYDALYANKASLVAINTIDA 540  
 Db 481 RLLRKDKLAPGTSOKYNNRTIDLDLTDVTILEKADGMSTYDALYANKASLVAINTIDA 540  
 QY 541 DNSGETLDERETAILDVLVHMGAYSKAEMLEKCRMMDJANGDKUDNFLEATRISDL 600  
 Db 541 DNSGETLDERETAILDVLVHMGAYSKAEMLEKCRMMDJANGDKUDNFLEATRISDL 600  
 QY 601 HRKEQODENIRRSTORPSVAKTATDPUVTLADKISKNTLVEHDIDPTCSEKVIDPK 660  
 Db 601 HRKEQODENIRRSTORPSVAKTATDPUVTLADKISKNTLVEHDIDPTCSEKVIDPK 660

OY	661	S	661	Db	361	SFKORMGIVESSALKELAVMRDRHDLDEBFKRYDKDGYTISISHCKVUNVTKLGL	420
Db	661	S	661	OY	479	PWMLRDKLAPGTDQSKVNNTFLDIDTVLAEADGMSYMDALVANKASLVAENII	538
AC	Q51D3;	DROME	PRELIMINARY;	PRT;	603	AA.	421
DT	01-FEB-2005	(TREMBLREL.	29, Created)				
DT	01-FEB-2005	(TREMBLREL.	29, Last sequence update)				
DE	RH46370P.	Name=rgC;					
GN	Drosophila melanogaster	(Fruit fly).					
OS	Eukaryota; Metazoa; Anthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Nematoda; Drosophilidae; Drosophila.						
OC							
OX							
RP							
RC							
RA	Stapleton M., Carlson J., Chavez C., Fries B., George R., Paclob J., Park S., Wan K., Yu C., Rubin G.M., Celinker S., Submitted (OCT-2004) to the EMBL/Genbank/DDBJ databases.			ID	Q1Q229	ANOGA	RESULT 2
RL	EMBL; BT01559; ANX36844.1; -; rRNA.			AC	070229	ANOGA	PRELIMINARY;
DR	GO; GO:0005509; F:calcium ion binding; IEA.			DT	01-MAR-2004	(TREMBLREL.	26, Created)
DR	GO; GO:0016787; P:hydrolase activity; IEA.			DT	01-MAR-2004	(TREMBLREL.	26, Last sequence update)
DR	InterPro; IPR002498; BP-hand.			DB	01-MAR-2004	(TREMBLREL.	26, Last annotation update)
DR	InterPro; IPR004943; M-Pesterase.			GN	ENSANG30000004336	(Fragment).	
DR	InterPro; IPR006186; T_Phtase_apAH.			ORFnames=ENSANG30000003837;			
DR	PFam; PF00036; efamily; 3.			OS	anophales_gambiae str. PEST.		
DR	PRINTS; PRO00114; STPHPTASE.			OC	Eukaryota; Metazoa; Anthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea; Culicidae;		
DR	PROSITE; PS00012; SER_THR_PHOSPHATASE; UNKNOWN 1.			OC	Anophelinae; Anopheles.		
DR	SEQUENCE; 603 AA; 65121 MW; A2FBC6A32CE2403 CRC64;			OK	NCBI_TAXID=180454;		
DR	Query Match 90.7%; Score 3121; DB 2; Length 603; Best Local Similarity 99.8%; Pred. No. 5.3e101; Matches 602; Conservative 0; Mismatches 1; Indels 0; Gaps 0;			RN	659	KKS 661	
OY	59 MPQAAGRENOYQOSAHVUSUDDKDQJTERGDVNAKEIPLRKWHDLLVPRKGNN 118 1 MPQAGREKNOYQOSAHVUSUDDKDQJTERGDVNAKEIPLRKWHDLLVPRKGNN 60	RC	NUCLEOTIDE SEQUENCE.	Db	601	KKS 603	
Db	119 RLHPPKVVALRERAKSAKSQLPNTPSPVASTAVSQCVTCGDLAGKQLDDLVVLRKNGLSS 178 61 RLHPPKVVALRERAKSAKSQLPNTPSPVASTAVSQCVTCGDLAGKQLDDLVVLRKNGLSS 120	RC	STRAIN=PEST;	Db	361	SFKORMGIVESSALKELAVMRDRHDLDEBFKRYDKDGYTISISHCKVUNVTKLGL	420
OY	179 SNPYVENGDFVDRGKRGKLEVLLILSLYLRPAVFLNRGNERDSVNYARGPIREYESK 238 121 SNPYVENGDFVDRGKRGKLEVLLILSLYLRPAVFLNRGNERDSVNYARGPIREYESK 180	RC	The Anopheles gambiae Sequence Committee; Submitted (APR-2004) to the EMBL/GenBank/DDBJ databases.	Db	479	PWMLRDKLAPGTDQSKVNNTFLDIDTVLAEADGMSYMDALVANKASLVAENII	538
Db	239 YPPNHKRILARIDEVYRMLPLGSVLSVNLVHGGSDSTDLSLIKSDIKRVSIRPP 298 181 YPPNHKRILARIDEVYRMLPLGSVLSVNLVHGGSDSTDLSLIKSDIKRVSIRPP 240	CC	-1- CAUTION: The sequence shown here is derived from an EMBL/GenBank/ DDBJ whole genome shotgun (WGS) entry which is preliminary data.	Db	481	DADNSGETLDPETATDLYVAMPGAYSKAEMLEKCRMDLNGDGKVUNEPLEARLIS	540
OY	299 LTGEPADKTEWQIFDMMWSDROATMGCVPNLTIGAGCWGRGDVDTNFLQRHLSTVIR 358 241 LTGEPADKTEWQIFDMMWSDROATMGCVPNLTIGAGCWGRGDVDTNFLQRHLSTVIR 300	CC	EMBL; ARAB100866; EAA12933.2; -; Genomic DNA.	Db	539	DADNSGETLDPETATDLYVAMPGAYSKAEMLEKCRMDLNGDGKVUNEPLEARLIS	598
Db	359 SHCKPNHGFMHDNKITIFSFASNYVYIGNSNGAYTRNLNOLMPHVOYSAASQTKRL 418 301 SHECKPNHGFMHDNKITIFSFASNYVYIGNSNGAYTRNLNOLMPHVOYSAASQTKRL 360	CC	DR; GO; GO:0016787; P:hydrolase activity; IEA.	Db	541	DLMRKBDENIRRSTGRPSVAKTADPVTLADKISKNTLVEHDIDPTCESKVIDP	600
OY	419 SPKORMGIVESSALKELAVMRDRHDLDEBFKRYDKDGYTISISHCKVUNVTKLGL 478	CC	DR; InterPro; IPR00048; BP-region.	Db			
DR	InterPro; IPR00483; M-Pesterase.			DR			
DR	InterPro; IPR006186; T_Phtase_apAH.			DR			
DR	PFam; PF00036; efamily; 2.			DR			
DR	PFam; PF00612; IQ; 1.			DR			
DR	PRINTS; PRO0014; STPHPTASE.			DR			
DR	PRODom; PRO00012; BP-hand; 1.			DR			
DR	PRODom; PRO00252; T_Phtase_apAH; 1.			DR			
DR	SMART; SM00054; EPh; 2.			DR			
DR	SMART; SM00056; PP2Ac; 1.			DR			
DR	PROSITE; PS00018; EQ_HAND; UNKNOWN 2.			DR			
DR	PROSITE; PS00016; IQ; 1.			DR			
DR	PROSITE; PS000125; SER_THR_PHOSPHATASE; UNKNOWN_1.			DR			
FT	NON_TER 1			FT			
FT	NON_TER 613			FT			
SO	SEQUENCE 613 AA; 69803 MW; BA0AARAFEP89D548 CRC64;			SO			
Query Match 56.1%; Score 1930.5; DB 2; Length 613; Matches 366; Conservative 101; Mismatches 130; Indels 17; Gaps 6;							



ID	QASQBL_TETIG PRELIMINARY;	PRT;	705 AA.	Db	357 RHLHSQQLSGQVNWSVVEELKKRRLLAGFDOSYGBPQRSDSDSPGEGATEADBEHWK	416
AC	QISQBL;			Qy	312 QIFDIMSPPDQATMGCTPNLTLRGAGWPGPDYTDNPLQRHRISVIRSHHECKENGHPHM	371
DT	13-SEP-2005 (Tremblrel: 31, Last sequence update)			Db	417 QTWDILSPDMFQSGCIPNEVRGGCTWGPDVTEKVARDHQOLQYVSEHECKQBGPERCH	476
DT	13-SEP-2005 (Tremblrel: 31, Last annotation update)			Qy	372 DNKLITTSFSASNYYAIGSNKGAYIRLNLQMLPHFQVQISAASQTKRISFKORGIVESSA	431
DB	Chromosome 4 SCAF1453, whole genome shotgun sequence (Fragment).			Db	477 DRRVLITPSASNYIEVSNSNRGAVIRMDLIPHFWQY-QASKTCRELTILQSVGRTERSA	535
GN	ORIGIN=GSTENG0001444001; OS=Tetraodon nigroviridis (Green puffer)			Qy	432 LKEBLAVNRDRDELEBFRKCDPKGSYISSHWCYMENTKLGAPWLRDCLKPT	491
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Buteleostomi; Actinopterygii; Neopterygii; Teleostei; Tetraodontiformes; Acanthomorpha; Actanopterigii; Percormorpha; Tetraodontidae; Tetradontoidae; Tetraodontidae; Tetradon.			Db	536 LRALKEGDFVHKSDLMASFEDPNNTGMLSLRHMASATERVNLGLPWRVLPOLVST	595
OC	[1]; NCBI_TAXID=99883;			Qy	492 DSQKVNNTL-DLDITVILERADOMSVDALYANKSVAIPNIDAONSGETLDE	550
OX	RAJallon O., Aury J.M., Brunet F., Petit J.L., Strange-Thomann N., Maucler E., Bouneau I., Fischer C., Ozouf-Costaz C., Bernot A., Niclaud S., Jaffe D., Fisher S., Dutfalla G., Bousat C., Seguens B., Basilia C., Salamoubat M., Levy M., Boudet N., Castellano S., Anthouard V., Jubin C., Castelli V., Katinka M., Vacherie B., Bleumont C., Skalli Z., Cattolico L., Poulaen J., De Bernardini V., Cruaud C., Duprat S., Brottier P., Coutanceau J.P., Gouzy J., Parra G., Iardier G., Chapple C., McFarlan P., Bosak S., Kellis M., Wolff J.N., Guigo R., Zody M.C., Mesirov J., Lindblad-Toh K., Birren B., Nusbaum C., Kahn D., Robinson-Rechavi M., Laudet V., Schachter V., Quétier F., Saunier M., Sarpalić C., Wincker P., Lander E.S., Weissenbach J., Roest Crollius H.; reveals genome duplication in the teleost fish "Tetraodon nigroviridis reveals the early vertebrate proto-karyotype"; Nature 431:946-957 (2004).			Db	596 ESGMWDYDQOWINLFRSTRPEKLEM-ADN-SILETMWRYNHNLETPRIDRTHGSLSPB	653
RN	RP	NUCLEOTIDE SEQUENCE.		Qy	551 FBTAAIDLVAHMPGAYKRAEMLKCRMADLNQDGKVUNLFNELEAFRISDL	600
RG	Genoscope; Whitehead Institute Centre for Genome Research; submitted (EBI-GenBank/DDBJ) to the EMBL/GenBank/DDBJ databases.			Db	654 FHTQWKLUSSHKMEISDKAVTDLAQSIDFNKQGSDISIDINEFWEARFLVH	703
RL	RA	- - CAUTION: The sequence shown here is derived from an EMBL/GenBank/DDBJ whole genome shotgun (WGS) entry which is preliminary data.		DT	16-OCT-2001 (Rel. 40, Last sequence update)	
CC	RA	EMBL; CAA01014533; CAR97171_1; -; Genomic_DNA.		DT	13-SEP-2005 (Rel. 48, Last sequence update)	
CC	RA	NON_TER 1		DB	Serine/threonine protein phosphatase with EF-hands-2 (EC 3.1.3.16) (PPBP-2)	
FT	RA	SEQUENCE 705 NON_TER 705		DB	Name=PPBP-2;	
FT	RA	705 AA; 81541 MW; E4513DBA03B1POA CRC64;		GN	OS	
FT	RA	NUCLEOTIDE SEQUENCE.		GN	Mus musculus (Mouse).	
FT	RA	Best Local Similarity 39.8%; Score 1371.5; DB 2; Length 705; Matches 278; Conservative 117; Mismatches 186; Indels 129; Gaps 10;		GN	OC	
Qy	5 AIRBALIFQWYRQHARREMRQWQKFQNLEYASQDQALYKPFENDLKHMPDAG	64		GN	OC	
Db	9 AIRRALIQTQWYRQHARREMRQWQKFQNLEYASQDQALYKPFENDLKHMPDAG	68		GN	OC	
Qy	65 RGY-----			GN	OC	
Db	69 ERNLISHFRKNEUCGRADWERYFESYSSVERVPYS-----			GN	OC	
Qy	107 LLDFVFKERGNRLHPKVYLAKRBAKSLSKOLPNISPVSTAVSQOQVUGLHGKDDL	166		GN	OC	
Db	119 KLYVEAFKHQQ-OHARTVQLOLGETWFLRILPNINGVSACONREITICGDGHLDL	176		GN	OC	
Qy	157 LWVUHK-----NGLPSSSNPYPNGFUDRKREGILVLLISLXMAPAVPLARGN	219		GN	OC	
Db	177 LVVYKRELPVQAGQFPSSBKPYVNGFDVDRGKSSTILLFLFGPLVYPDVHLRGN	236		GN	OC	
Qy	220 HEQSMWARYGPFRVESVKPRMHKILAPIDEVRYWPLGSVLSNRSVLIWGGPSDSTS	279		GN	OC	
Db	237 HEHVIVNLYGFTKEWKGKTYRNGKIKLQLQKFSMPLATDHDKLIVWGGISPTD	296		GN	OC	
Qy	280 LDLTISDQGKYSILRPLP-----			GN	OC	
Db	297 LDAJARVDRHKYVSLRPPKLNQQAANEKragaaGPMEGRRVYTLKSTADHSHKRP	356		GN	OC	
Qy	300 -----			GN	OC	
RP	NUCLEOTIDE SEQUENCE.			Qy	312 QIFDIMSPPDQATMGCTPNLTLRGAGWPGPDYTDNPLQRHRISVIRSHHECKENGHPHM	371
RP	NUCLEOTIDE SEQUENCE.			Db	417 QTWDILSPDMFQSGCIPNEVRGGCTWGPDVTEKVARDHQOLQYVSEHECKQBGPERCH	476
RC	RC	TISSUE=Eye; MEDLINE=9741020; PubMed=9326663; DOI=10.1073/pnas.94.21.11639; Proc. Natl. Acad. Sci. U.S.A. 94:11639-11644(1997).		Qy	372 DNKLITTSFSASNYYAIGSNKGAYIRLNLQMLPHFQVQISAASQTKRISFKORGIVESSA	431
RC	RC	NUCLEOTIDE SEQUENCE [LARGE SCALE MEGA].		Db	477 DRRVLITPSASNYIEVSNSNRGAVIRMDLIPHFWQY-QASKTCRELTILQSVGRTERSA	535
RC	RC	STRAIN=C57BL/6; TISSUE=Retina; MEDLINE=2238825; PubMed=24749732; DOI=10.1073/pnas.242603899; Strauberg R.L., Feingold E.A., Grouse L.H., Berge J.G., Schuler G.D., Klausner R.D., Colling R.S., Wagner L., Shemesh C.M., Schueler G.D., Altshuler S.P., Zeeberg B., Butow K.H., Schaffer C.P., Bhat N.K., Hopkins R.P., Jordan H., Moore T., Max S.I., Wang J., Heilek F., Diatchenko L., Marusina K., Farmer A.R., Rubin G.M., Hong L., Staption M., Soarees M.B., Bonaldo M.F., Casavant T.L., Schatz T.E., Brownstein M.J., Usdin T.B., Toshimuki S., Carninci P., Prange C., Raha S.S., Logquastano N.A., Peters J.W., Abramson R.D., Mullany S.J., Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H., Richards S., Woolley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W., Villalon D.K., Munro D.M., Sodergren E.J., Lu X., Gibbs R.A., Fahey J., Heiton E., Kettman M., Madan A., Rodriguez S., Sanchez A., Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G., Blakely R.W., Touchman J.W., Green E.D., Dickson M.C., Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E., Schnecko A., Schein J.E., Jones S.J.M., Maira M.A., RT		Qy	432 LKEBLAVNRDRDELEBFRKCDPKGSYISSHWCYMENTKLGAPWLRDCLKPT	491
RC	RC	"Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences"; Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002); I- FUNCTION: May play a role in phototransduction. May		Db	536 LRALKEGDFVHKSDLMASFEDPNNTGMLSLRHMASATERVNLGLPWRVLPOLVST	595



DR PF00149; Metallophos; 1.  
 DR PIRSPF; PIRSP00912; PPEPF; 1.  
 DR PRINTS; PRO0114; STPHTPTASE  
 DR PRODOM; PD000112; EF-hand; 1.  
 DR PRODOM; PD000252; T\_phthase\_apah; 1.  
 DR SMART; SM00154; EFh; 2.  
 DR PROSITE; PS00018; BP\_HAND; UNKNOWN<sup>2</sup>; UNKNOWN\_1.  
 KW Hypothetical protein.  
 SQ SEQUENCE 707 AA; 79840 MW; 4ADE2CA1D186C630 CRC64;  
 Query Match 39.0%; Score 1341; DB 2; Length 707;  
 Best Local Similarity 42.4%; Pred. No. 7, 3e-73; Matches 269; Conservative 118; Mismatches 192; Indels 56; Gaps 12;  
 Matches 269; Conservative 118; Mismatches 192; Indels 56; Gaps 12;  
 QY 6 TRAIFIQKWRKQHARREMRMRCNWOIIFONLBYASASDQABLYKFENDLIRKMPDQAGR 65  
 DB 83 IKSALLIQKWRKQHARREMRMRCNWOIIFONLBYASASDQABLYKFENDLIRKMPDQAGR 65  
 QY 103 NHIDLLIDVFRKKRGRGNRLHPIKYVALLIREAKAASKLQLPNISPVSTAVSQVTVCGLDNGK 162  
 DB 203 PQVAKMTEAF-KVNUKVLPKVYLMLHEARKIFPKMSPVSRISTISNQITICGJHGK 260  
 QY 163 LDDILVVLWHLKNGLPPSSNPYVNGDFVDRKGKLEVVLLLSLYLAEPNAFLNQRGHED 222  
 DB 261 FPDCLLTYKNGLPPSVDPNYPVNGDFVDRKGKLEVVLLLSLYLAEPNAFLNQRGHED 222  
 QY 320 223 SYMVARYGPIREVESKPYKPRHHKELAPIDEVWRMLPLGSVLSNRVLIHGGSDSISLDL 282  
 DB 321 HMTNLRYGPIKELSTKVDLSPTIRLIEDVFSWLPATIVDVKDIFVWNGGSDPTEVK 380  
 QY 381 LDKIPRHRHQFQSVLRPVPGKMDSAKERNQAVSDEWKRQMDIMWSDFQNKCGCWPNVFRG 440  
 QY 335 AGSWFGEPTDIFLORHLVSYVIRSHACKPNSHTEFMNDKITTFSASNNYALGSNKGAY 394  
 DB 441 GSYVFGIADITASPLRKIGFRLVLRSHCKPNSHTEFMNDKITTFSASNNYALGSNKGAY 500  
 QY 395 IRANOLQMPHQYQYISAASQTKRLS-FKORMGIVESSALKELAVRMRDHRRDHLBDBERRY 453  
 DB 501 VRFIGKSKPHFQYMM- ASKTRKSTLURGLVRSRVAKEKLKQLSSFHTELOKEENV 558  
 QY 512 454 DPKDGGTISIHWCKMHEVNTVKGFLPWLRLDKLAPGTSOK-VNMRTRLDLDTYDILB 512  
 DB 559 DLERKSGNIPILKWSECVRIGLNLPWITALPKVATLSDDGKFWYKEDRKIAQVGTHA 618  
 QY 513 ARADGMSWMDAYANKASLVAENIIDADSBSBTIDPFTETIDLUYAHMPGAYSK--- 568  
 DB 619 QEQD--I-IVESLXHRSKSTLETFRMKDQNSGCVSKMKEPIDACRVL---GKYTKRPLQ 670  
 QY 569 ---AEMLAKCRMADDLNGKUDLNEFRAFRSLD 599  
 DB 671 TDYIQAQAES--IDPNKDFP1DLNRLKEAFLVD 702  
 RESULT 8  
 PPE2\_HUMAN STANDARD PRT; 753 AA.  
 ID PPE2\_HUMAN  
 AC 014830; OL4831;  
 DT 16-Oct-2001 (Rel. 40, Created)  
 DT 16-Oct-2001 (Rel. 40, Last sequence update)  
 DT 13-SEP-2005 (Rel. 48, Last annotation update)  
 DE Serine/threonine protein phosphatase with EF-hands-2 (EC 3.1.3.16)  
 DE (PPEPF-2).  
 Name=PPEPF-2;  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Homo.;  
 OC NCBI\_TaxID=9606;  
 RN [1]  
 RP NUCLEOTIDE\_SEQUENCE (ISOFORMS PPEPF-2(S) AND PPEPF-2(L)).  
 RX TISSUE-Retina; PubMed=9326633; DOI=10.1073/pnas.94.21.11639;  
 RA Sherman P.M., Sun H., Macke J.P., Williams J., Smallwood P.M.,  
 RA Nathans J.;  
 RT "Identification and characterization of a conserved family of protein  
 RT serine/threonine phosphatases homologous to Drosophila retinal  
 RT degeneration C."  
 RL Proc. Natl. Acad. Sci. U.S.A. 94:11639-11644 (1997).  
 CC -I- FUNCTION: May play a role in phototransduction. May  
 CC dephosphorylate photoactivated rhodopsin. May function as a  
 CC calcium sensing regulator of ionic currents, energy production or  
 CC synaptic transmission.  
 CC -I- CATALYTIC ACTIVITY: A phosphoprotein + H(2)O = a protein +  
 CC phosphate.  
 CC -I- COFACTOR: Binds 1 iron ion per subunit (By similarity).  
 CC -I- COFACTOR: Binds 1 manganese ion per subunit (By similarity).  
 CC -I- ENZYME REGULATION: Activated by calcium (By similarity).  
 CC -I- SUBCELLULAR LOCATION: Localized to photoreceptors. PPEPF-2(L) is at  
 CC least 2 fold more abundant in rod inner segments than in the outer  
 CC segments.  
 CC -I- ALTERNATIVE PRODUCTS:  
 CC Event=Alternative splicing; Named isoforms=2;  
 CC Name=PPEPF-2(L);  
 CC IsoID=O14830-1; SequenceDisplayed;  
 CC Name=PPEPF-2(S);  
 CC IsoID=O14830-2; Sequence=VSP 005103; VSP 005104;  
 CC -I- TISSUE SPECIFICITY: Retinal specific (Probable).  
 CC -I- SIMILARITY: Belongs to the PPE phosphatase family.  
 CC -I- SIMILARITY: Contains 1 IQ domain.  
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use as long as its current state is in no way modified and this statement is not  
 CC removed.  
 CC -----  
 DR EMBL; AF024456; AAB82796.1; - mRNA.  
 DR RISP; P09860; IDTL.  
 DR Ensembl; ENSG00000156194; Homo sapiens.  
 DR HGNC; HGNC:9244; PPEPF-2.  
 DR MM; 602256; -.  
 DR GO; GO:000472; P:protein serine/threonine phosphatase activity; TAS.  
 DR GO; GO:0005470; P:protein amino acid dephosphorylation; TAS.  
 DR InterPro; IPR01192; EF-Hand-type.  
 DR InterPro; IPR00248; EF-hand\_Ca<sup>2+</sup>-region.  
 DR InterPro; IPR00044; IQ\_Cam\_Bd-region.  
 DR InterPro; IPR004843; M-pesterate.  
 DR InterPro; IPR012008; PPEPF.  
 DR InterPro; IPR006186; T\_phthase\_apah.  
 DR PIRam; PF00036; erhand; 3.  
 DR PIRam; PF00036; erhand; 3.  
 DR PIRam; PF00044; Metallophos; 1.  
 DR PIRSP; PIR000912; PPEPF; 1.  
 DR PRINTS; PR00114; STPHTPTASE.  
 DR ProDom; PD000012; EF-hand; 1.  
 DR ProDom; PD000252; T\_phthase\_apah; 1.  
 DR SMART; SM0054; EFh; 3.  
 DR SMART; SM0015; IQ; 1.  
 DR SMART; SM00156; PP2Ac; 1.  
 DR PROSITE; PS00018; BP\_HAND; 1.  
 DR PROSITE; PS00096; IQ; FALSE\_NEG.  
 DR PROSITE; PS00125; SER\_THR\_PHOSPHATASE; 1.  
 KW Alternative splicing; Calcium; Hydrolase; Iron; Manganese;  
 KW Metal-binding; Polymerase; Protein phosphatase; Repeat;  
 KW Sensory transduction; Vision.



	SQ	SEQUENCE	653 AA:	75406 MW; A54FGR236554516P CRC64;
QY	6	IRALIPIQKWRHQARREMRQQCNWQIFQNLSEASBODQABLYKFENDLKHMPQAAGR	65	. Query Match 37.3%; Score 1285.5; DB 2; Length 653;
Db	83	IKSAILIQKWRCEARLEARRAATWQFPLAIVAGEBDQDQLKXDPPADVIRAMAENGK	142	Best Local Similarity 40.8%; Pred. No. 1.5e-69; Matches 261; Conservative 122; Mismatches 209; Indels 47; Gaps 9;
QY	66	KNOYQG-SAHVSYLDD--KDDLVEEFDINWAKIE-----	143	Matches 261; Conservatve 122; Mismatches 209; Indels 47; Gaps 9;
Db	143	GGYENGRNSPPLMSALSHYAKPSLMSEGSTVKOMLEDTSPTNVDIDRNYKGPTLSLPLDK	202	143 NHTDLIIVYPRKURGNGRNHAKPVYALLRBAKLIKQLENISPVSTAVSQVTUGDGHK
QY	103	:-----	261	103 NHDTDLIIVYPRKURGNGRNHAKPVYALLRBAKLIKQLENISPVSTAVSQVTUGDGHK
Db	203	PQVAKMIRAF--KVNVLKPKYUMLHARKTFKAMPVSRSITISNOVTCGDK	260	203 PQVAKMIRAF--KVNVLKPKYUMLHARKTFKAMPVSRSITISNOVTCGDK
QY	163	LDDLVVUHNGKLSSSPYVENGDFVDRGKRGLEVLILLSLXLAFFNAVEANGHEED	222	163 LDDLVVUHNGKLSSSPYVENGDFVDRGKRGLEVLILLSLXLAFFNAVEANGHEED
Db	261	FDDCILYKNGKPSVDPYPIFGDFVURGQGSIBEVLCUTLALVIVDPMSTIVLRGNHED	320	261 FDDCILYKNGKPSVDPYPIFGDFVURGQGSIBEVLCUTLALVIVDPMSTIVLRGNHED
QY	223	SUMMARYSPIREVESKYPRNHKILAPIDBRYVWPLGLSVLSRVLVPHGGSSTSDELK	282	223 SUMMARYSPIREVESKYPRNHKILAPIDBRYVWPLGLSVLSRVLVPHGGSSTSDELK
Db	321	HIMLRLYCPKTELSTKYKDSTPITRLLEDFVSWLPLATITIDRIVVHGIGSDOTENVSK	380	321 HIMLRLYCPKTELSTKYKDSTPITRLLEDFVSWLPLATITIDRIVVHGIGSDOTENVSK
QY	336	GWVFGPDYTDNFQHRHSYVIRSHECKENGHFMHDNKLTITFSASNYYAIGSNKAYI	395	336 GWVFGPDYTDNFQHRHSYVIRSHECKENGHFMHDNKLTITFSASNYYAIGSNKAYI
Db	441	GSYFAGADITASALEKHGRLUVSHECKEGFGEVSPHNTCLTFSASNYYETOSNRGAYV	500	441 GSYFAGADITASALEKHGRLUVSHECKEGFGEVSPHNTCLTFSASNYYETOSNRGAYV
QY	396	R-LNNOLMPHVFVYISAQSOTKRLS-FKORMGIVESSAKELAVERMDHRSLEDEBKRY	453	396 R-LNNOLMPHVFVYISAQSOTKRLS-FKORMGIVESSAKELAVERMDHRSLEDEBKRY
Db	501	KF1GSKSKOPHPVOM-ASKTRKKSSTERBLGVVEESAALKERKJSSPFHDLQKEPRIM	558	501 KF1GSKSKOPHPVOM-ASKTRKKSSTERBLGVVEESAALKERKJSSPFHDLQKEPRIM
QY	454	DPKDGYVISISHCKMVENVTKGPLWRLDKLAP-GTDSQKVNNTDILDTDVILE	512	454 DPKDGYVISISHCKMVENVTKGPLWRLDKLAP-GTDSQKVNNTDILDTDVILE
Db	559	DIEKSGKUPLKWDVCERTIGLNLPWTLAPKVATLSEGDGVYVMEYKEDRRAQVGTHA	618	559 DIEKSGKUPLKWDVCERTIGLNLPWTLAPKVATLSEGDGVYVMEYKEDRRAQVGTHA
QY	513	AEDGMGSYMDALYANKSLSVAIFIADNSGTTDEPFTAIDLVAHMPGAYSKARM-	571	513 AEDGMGSYMDALYANKSLSVAIFIADNSGTTDEPFTAIDLVAHMPGAYSKARM-
Db	619	QEAD--IYVSLYRHKSTLETLFRPMOKNSQCVSMKEDFACEVLL--GKTKRPLQ	670	619 QEAD--IYVSLYRHKSTLETLFRPMOKNSQCVSMKEDFACEVLL--GKTKRPLQ
QY	572	--LEKRMMDLNGDKVLDNEFLLEAFRLVD	599	572 --LEKRMMDLNGDKVLDNEFLLEAFRLVD
Db	671	TDVISQIARSIDFNKOGFIDLNBLLEAFRLVD	702	671 TDVISQIARSIDFNKOGFIDLNBLLEAFRLVD
<b>RESULT 10</b>				
Q4R4Y0	MACFA			
ID	O4R4Y0; MACFA PRELIMINARY;	PRT:	653 AA.	
AC	O4R4Y0;			
DT	13-SEP-2005 (TREMBLREL. 31, Created)			
DT	13-SEP-2005 (TREMBLREL. 31, Last sequence update)			
DB	Brain cDNA, clone: QTA-11374, similar to human protein phosphatase, BP hand calcium-binding domain 1 (PP1), transcript variant 1.			
DR	Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey, Macaca fascicularis (Crab eating macaque))			
OS	Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Buteleotomi; Mammalia; Buteraria; Buarchontoglires; Primates; Catarrhini; Cercopithecidae; Cercopithecinae; Macaca.			
OC	NCBI_TAXID=9541;			
OX	[1]			
RN	NUCLEOTIDE SEQUENCE.			
RR	International consortium for macaque cDNA sequencing, analysis;			
RA	"DNA sequences of macaque genes expressed in brain or testis and its evolutionary implications"; Comparative analyses between human and cynomolgus monkey cDNAs"; Submitted (JUN-2005) to the EMBL/GenBank/DDBJ databases.			
RR	[2]			
RR	NUCLEOTIDE SEQUENCE.			
RA	Osada N., Hirata M., Tanuma R., Kusuda J., Hida M., Suzuki Y., Sugano S., Gojobori T., Shen J.C.-K., Wu C.I., Hasimoto K.; "Substitution rate and structural divergence of 5'UTR evolution: Comparative analysis between human and cynomolgus monkey cDNAs"; Submitted (MAR-004) to the EMBL/GenBank/DDBJ databases.			
DR	EMBL; ABI69764; BAB01845.1; -; mRNA.			
<b>RESULT 11</b>				
RR	PEPL_HUMAN	STANDARD;	PRT:	653 AA.
ID	PEPL_HUMAN			
AC	O14829; O15253; Q9NU1; Q9UJH0;			
DT	16-OCT-2001 (Rel. 40, Crested)			
DT	16-OCT-2001 (Rel. 40, Last sequence update)			
DT	13-SEP-2005 (Rel. 48, Last annotation update)			
DB	Serine/threonine protein phosphatase with ER-hands-1 (EC 3.1.3.16)			
DB	(PPBP-1) (Protein phosphatase with ER calcium-binding domain) (PPBP)			
DB	(Serine/threonine protein phosphatase 7) (PP7).			
GN	Name=PPBP1; Synonyms=PPBP, PP7C;			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Buteleotomi; Mammalia; Buteraria; Buarchontoglires; Primates; Catarrhini; Homiidae;			
OC	Homo.			
NCBI_TAXID=9606;				
[1]				
RR	NUCLEOTIDE SEQUENCE.			
RC	TISSUE=Retina;			
RX	MEDLINE=97471020; PubMed=9326663; DOI=10.1073/pnas.94.21.11639;			
RA	Sherman P.M., Sun H., Macke J.P., Williams J., Smallwood P.M., Nathans J.;			

- RT "Identification and characterization of a conserved family of protein  
RT serine/threonine phosphatases homologous to *Drosophila* retinal  
RT degeneration C";  
RT Proc. Natl. Acad. Sci. U.S.A. 94:11639-11644 (1997).  
RN [2]  
RN NUCLEOTIDE SEQUENCE, AND ALTERNATIVE SPLICING.  
RC TISSUE=Fetal brain;  
RX MEDLINE=9758589; PubMed=921565; DOI=10.1093/hmg/6.7.1137;  
RA Montini E., Ruggeri R.I., van de Vosse S., Adolfi G., Mariani M.,  
RA Puccia A.A., Consales G.G., den Dunnen J.T., Ballabio A., Franco B.;  
RT "A novel human serine-threonine phosphatase related to the *Drosophila*  
RT retinal degeneration C (*rdcG*) gene is selectively expressed in sensory  
neurons of neural crest origin.";  
RL Hum. Mol. Genet. 6:1137-1145 (1997).  
RN [3]  
RN NUCLEOTIDE SEQUENCE.  
RC TISSUE=Retina;  
RX MEDLINE=94104127; PubMed=9430693; DOI=10.1074/jbc.273.3.1462;  
RA Huang X., Honkanen R.E.;  
RT "Molecular cloning, expression, and characterization of a novel human  
RT serine/threonine protein phosphatase, PP7, that is homologous to  
RT *Drosophila* retinal degeneration C gene product (*rdcG*).";  
RL J. Biol. Chem. 273:1462-1468 (1998).  
RN [4]  
RN NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].  
RX PubMed=15772651; DOI=10.1038/nature03401;  
RA Ross M.T., Grafham D.V., Coffey A.J., Scherer S., McLay K., Muzny D.,  
RA Platzer M., Howell G.R., Burrows C., Bird C.P., Frankish A.,  
RA Lovell P.L., Howe K.L., Asturte J.D., Fulton R.S., Sudrak R., Wen G.,  
RA Jones M.C., Hurles M.E., Andrews T.D., Scott C.E., Searle S.,  
RA Ramser J., Whittaker A., Deadman R., Carter N.P., Hunt S.E., Chen R.,  
RA Cree A., Gunnarsson P., Havlik P., Hodgson A., Metzker M.L.,  
RA Richards S., Scott G., Stoffelen P., Sodegran B., Wheeler D.A.,  
RA Worley K.C., Ainscough R., Ambrose K.D., Ansari-Lari M.A., Aradhya S.,  
RA Ashwell R.I., Babbage A.K., Baguley C.L., Ballabio A., Banerjee R.,  
RA Barker G.E., Barlow K.F., Barrett I.P., Bates K.N., Beare D.M.,  
RA Beasley H., Beasley O., Beck A., Bethel G., Blechschmidt K., Brady N.,  
RA Bray-Alien S., Bridgeman A.M., Brown A.J., Brown M.J., Bonnin D.,  
RA Burton J., Bye J.M., Garder C., Garrel L., Chako J., Chapman J.C.,  
RA Chavez D., Chen E., Chen G., Chen Y., Chen Z., Chinault C.,  
RA Ciccodicola A., Clark S.Y., Clarke C.M., Clegg S.,  
RA Clerc-Blankenburg K., Clifford K., Cobley V., Cole C.G., Conquer J.S.,  
RA Corby N., Connor R.B., David R., Davies J., Davis C., Davis J.,  
RA Delgado O., Desbazo D., Dhami P., Ding Y., Dinh H., Dodswork H.,  
RA Draper H., Dugan-Rocha S., Dunham A., Dunn M., Durbin K.J., Dutta I.,  
RA Badde T., Ellwood M., Emery-Cohen A., Errington H., Evans K.L.,  
RA Paulk L., Francis P., Frankland J., Fraser A.B., Galgoczy P.,  
RA Gilbert J., Gill R., Gloeckner G., Gregory S.G., Grubbs S.,  
RA Griffiths C., Grocock R., Gu Y., Gwilliam R., Hamilton C., Hart E.A.,  
RA Hawes A., Heath P.D., Heitmann K., Henning K., Hernández J.,  
RA Hinzmann B., Ho S., Hoffe M., Howden P.J., Huckle B.J., Hume J.,  
RA Hunt P.J., Hunt A.R., Isherwood J., Jacob L., Johnson D., Jones S.,  
RA Hunt P.J., Joseph S., Keehan S., Kershaw J.K., Khan Z.,  
RA Kloschis P., Klages S., Knights A.J., Kosiura A., Kovar-Smith C.,  
RA Laird G.K., Landorf C., Lawlor S., Leverenza M., Lewis L., Liu W.,  
RA Lloyd C., Lloyd D.M., Louie H., Loveland J.E., Lovell J.D.,  
RA Lozado R., Lu J., Lyne R., Ma J., Maheshwari M., Matthews L.H.,  
RA McDowell J., McLaren S., McMurray A., Meidl P., Maitinger T.,  
RA Milne J., Miner G., Mistri Y., Morgan M., Morris J., Mueller I.,  
RA Mullikin J.C., Nguyen N., Nordstek G., Nyakatura G., O'dell C.N.,  
RA Okwouo G., Palmer S., Pandian R., Parker D., Parrish J.,  
RA Pasternak S., Patel D., Pearce A.V., Pearson D.M., Pelan S.B.,  
RA Perez L., Porter K.M., Ramsey Y., Reichwald K., Rhodes S.,  
RA Ridder K.A., Schlessinger D., Schueler M.G., Sehra H.K.,  
RA Shaw-Smith C., Shen E.M., Showkneen R., Skuce C.D.,  
RA Smith M.L., Soheran B.C., Steingrubler H.B., Steward C.A., Storey R.,  
RA Swann R.M., Swarbrick D., Tabor P.E., Taudien S., Taylor T.,  
RA Teague B., Thomas K., Thorpe A., Timms Tracey A., Trevanian S.,  
RA Trojans A.C., d'Ursi M., Verduzzo D., Vilasana D., Waldron L.,  
RA Wall M., Wang Q., Warren J., Warry G.L., Wei X., West A.,  
RA Whitehead S.L., Whitley M.N., Wilkison J.E., Willey D.L.,  
RA Williams G., Williams L., Williamson A., Williamson H., Wimling L.,
- RA Woodmansey R.L., Wray P.W., Yen J., Zhang J., Zhou J., Zoghlbi H.,  
RA Zorilla S., Buck D., Reinhardt R., Poutak A., Rosenthal A.,  
RA Lehrach H., Meinild A., Minx P.J., Hillier L.W., Willard H.F.,  
RA Wilson R.K., Waterton R.H., Rice C.M., Vaudin M., Coulson A.,  
RA Nelson D.L., Weinstock G., Sulston J.E., Durbin R., Hubbard T.,  
RA Gibbs R.A., Beck S., Rogers J., Bentley D.R.,  
RA "The DNA sequence of the human X chromosome.,";  
RN Nature 434:325-337 (2005).  
RN [5]  
RN NUCLEOTIDE SEQUENCE [LARGE SCALE RNA] (ISOFORM 1).  
RC TISSUE=Testicis;  
RX MEDLINE=2238827; PubMed=12477932; DOI=10.1073/pnas.242603899;  
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins P.S., Wagner L., Shevchenko C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Bustow K.H., Schefer C.F., Bhat N.K.,  
RA Hopkins R.P., Jordon H., Moore T., Max S.I., Wang J., Heien F.,  
RA Diatchenko L., Marsina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Usdin T.B., Toshiyuki I., Carninci P., Prange C.,  
RA Raha S.S., Loquaiello N.A., Peters G.J., Abramson R.D., Mullahay S.J.,  
RA Rodriguez A.C., Grimwood J., Schmutz J.J., Myers R.M.,  
RA Bosak S.A., McKernan K.J., Malek J.A., Gunnarsson P.H.,  
RA Richards S., Worley K.P., Skalska U., Smalius D.E.,  
RA Villon D.K., Muzny D.M., Sodegran B.J., Lu X., Gibbs R.A.,  
RA Fahy J., Helton E., Ketteman M., Madan N., Rodrigues S., Sanchez A.,  
RA Whiting M., Madan A.C., Young A.C., Shvechnikov Y., Bouffard G.G.,  
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalius D.E.,  
RA Schnarch A., Schein J.B., Jones S.J.M., Marra M.A.,  
RA "Generation and initial analysis of more than 15,000 full-length human  
RT and mouse cDNA sequences.,";  
RT Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).  
RN [6]  
RN SPlice Isoform(s) That Are Potential NMD Target(s).  
RX PubMed=14759259; DOI=10.1186/gb-2004-5-2-r8;  
RA Hillman R.T., Green R.B., Brenner S.E.,  
RT "An unappreciated role for RNA surveillance.,";  
RT Genome Biol. 5:RESEARCH008.1-RESEARCH008.16 (2004).  
CC -I- FUNCTION: May have a role in the recovery or adaptation response  
of photoreceptors. May have a role in development.  
CC -I- CATALYTIC ACTIVITY: A phosphoprotein + H<sub>2</sub>O = a protein +  
phosphate.  
CC -I- COFACTOR: Binds 1 iron ion per subunit (By similarity).  
CC -I- COFACTOR: Binds 1 manganese ion per subunit (By similarity).  
CC -I- COPACTOR: Magnesium (By similarity).  
CC -I- ENZYME REGULATION: Activated by calcium.  
CC -I- BIOPHYSICOCHEMICAL PROPERTIES:  
CC Optimum pH is 8.0;  
CC -I- ALTERNATIVE PRODUCTS:  
CC Event=Alternative Splicing; Named isoforms=5;  
CC Name=1;  
CC IsoID=014829-1; Sequence=Displayed;  
CC Name=1A;  
CC IsoID=014829-2; Sequence=VSP\_005098;  
CC Name=1B;  
CC IsoID=014829-3; Sequence=VSP\_005099;  
CC Name=2;  
CC IsoID=014829-4; Sequence=VSP\_005100; VSP\_005101;  
CC Notes=May be produced at very low levels due to a premature stop  
codon in the mRNA, leading to nonsense-mediated mRNA decay;  
CC Name=3;  
CC IsoID=014829-5; Sequence=VSP\_005102;  
CC Notes=May have no functional significance;  
CC -I- TISSUE SPECIFICITY: Detected in retina and retinal derived Y-79  
retinoblastoma cells. Also found in fetal brain.  
CC -I- SIMILARITY: Belongs to the PPP phosphatase family.  
CC -I- SIMILARITY: Contains 3 EF-hand domains.  
CC -I- SIMILARITY: Contains 1 IQ domain.  
CC This Swiss-Prot entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its



[2] RN  
 NUCLEOTIDE SEQUENCE.  
 RG Institute Centre for Genome Research;  
 Genoscope; Whitehead Institute for Biotechnology;  
 RLU Submitter (FEB-2004) to the EMBL/GenBank/DBJ databases.  
 CC -1 CAUTION: The sequence shown here is derived from an  
 EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is  
 preliminary data.  
 DR EMBL; CABO1014581; CAP99815.1; -, Genomic\_DNA.  
 FT NON\_TSR 1  
 FT 671 671  
 SQ SRQURE 671 AA; 77077 MN; 54AF1BF922697BD CRC64;

Query Match 36.0%; Score 1238.5; DB 2; Length 671;  
 Best Local Similarity 37.7%; Pred. No. 1..1e-66; Gaps 7;  
 Matches 254; Conservative 122; Mismatches 212; Indels 85; Gaps 7;

Db 304 --PLDKTIEWQQIDIMMSPDQATMGCPVNTRGAGWFGPDYTDNFLQRHLRSYVTRSH 360  
 Qy : |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||:  
 Db 372 PSPQHDEEWKQVWILSDPKLKGCPNTRGGCYFGPYVTRRLLQGLQLIRSH 431  
 Qy : |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||:  
 Db 361 BCKPENGHFMHDWKITIFSAASNYATGSNKAYIRLNQMLPHVOYISAQSOKRLSF 420  
 Qy : |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||:  
 Db 432 ECKIOBGEYLCHSGQVITIFSAASNYTEEGSNRGAYIKVGREMPFRYQY-QVSRSRKLT 490  
 Qy : |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||:  
 Db 421 KOMGMIVSSEALCILAVEMRDHDELEDFRKDPDKGSYIISWCKMENVTKGW 480  
 Qy : |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||:  
 Db 491 TQYRAAAGSAFRAKKEKLPTHSSELIGPOEDHNNTGTLVSEWAQVLFGLRDLPW 550  
 Qy : |||: |||: |||: |||: |||: |||: |||: |||: |||: |||:  
 Db 481 RLJDKLAPGTDSQVKNTNRDLDLTVILBARADGMSWMDLYANKASVLAIFIIDA 540  
 Qy : |||: |||: |||: |||: |||: |||: |||: |||: |||:  
 Db 551 DASGRITIDEFFERAILDVAMHPGASKARMLKCRMNDPGKVULNEFLAEFLSDL 600  
 Qy : |||: |||: |||: |||: |||: |||: |||: |||:  
 Db 609 DQGSLISIEEFRHTWHLFSAHLGKIKDNRAIDLRSIDFNKDGSDFNFELAEFRV-V 666  
 Qy : |||: |||: |||: |||:  
 Db 601 HRHQDQENIR 612  
 Qy : |||: |||:  
 Db 667 HKUDNKOQPNR 678  
 Qy : |||:  
 RESULT 13  
 ID Q4SHR7\_TRTNG  
 AC Q4SHR7  
 DT 13-SEP-2005 (TREMBlre. 31, Created)  
 DT 13-SEP-2005 (TREMBlre. 31, Last sequence update)  
 DT 13-SEP-2005 (TREMBlre. 31, Last annotation update)  
 DT Chromosome 5 SCAT14581, whole genome shotgun sequence.  
 (Fragment).  
 DB ORFNames=GSTENG00118033001,  
 GN Green puffer.  
 OS Tetraodon nigroviridis (Green puffer).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Buteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Neoteleostei;  
 OC Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;  
 OC Tetradontoidae; Tetraodontidae; Tetraodon.  
 OX NCBI\_TAXID=99883;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE.  
 RA Jallion O., Aury J.M., Brunet F., Petit J.I., Stange-Thomann N.,  
 RA Maucelle E., Bouneau L., Fischer C., Ozouf-Cozic C., Bernot A.,  
 RA Nicaud S., Jaffe D., Fisher S., Lutfalla G., Dobat C., Segurens B.,  
 RA Dasilva C., Salanoubat M., Levy M., Boudet N., Castellano S.,  
 RA Anthouard V., Jubin C., Castelli V., Katinka M., Vacherie B.,  
 RA Bilemont C., Skalli Z., Cattolico L., Poulain J., De Bernardis V.,  
 RA Cruaud C., Duprat S., Brottier P., Coutanceau J.P., Goury J.,  
 RA Parra G., Hardier G., Chapple C., McErwan P., Bosak S.,  
 RA Kellis M., Wolff J.N., Guigo R., Zody M.C., Medrano J., Lindblad-Toh K., Birren B., Nusbaum C., Kahn D., Robinson-Rochavi M.,  
 RA Laudet V., Schachter V., Oettler F., Saurin W., Scarpelli C., Wincker P., Lander E.S., Weissenbach J., Roest Crollius H.,  
 RT "Genome duplication in the teleost fish Tetraodon nigroviridis reveals the early vertebrate proto-karyotype";  
 RT Nature 431:946-957(2004).  
 RL

[2] RN  
 NUCLEOTIDE SEQUENCE.  
 RG Institute Centre for Genome Research;  
 Genoscope; Whitehead Institute for Biotechnology;  
 RLU Submitter (FEB-2004) to the EMBL/GenBank/DBJ databases.  
 CC -1 CAUTION: The sequence shown here is derived from an  
 EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is  
 preliminary data.  
 DR EMBL; CABO1014581; CAP99815.1; -, Genomic\_DNA.  
 FT NON\_TSR 1  
 FT 671 671  
 SQ SRQURE 671 AA; 77077 MN; 54AF1BF922697BD CRC64;

Query Match 36.0%; Score 1238.5; DB 2; Length 671;  
 Best Local Similarity 37.7%; Pred. No. 1..1e-66; Gaps 7;  
 Matches 254; Conservative 122; Mismatches 212; Indels 85; Gaps 7;

Db 1 AKAAVILVORWPRRYMMHEMRERRYWNIFOSIBEVASBODOLSSPFSEMLNNITHLNG 60  
 Qy 5 ATRAIAFTQKVRHKARREMRQRCWQIQTLEYASQDPAELKFENDLIKHMPCAG 64  
 Db 119 RLHARTVQLQYETRKLUKOMENVIALSTTYTKETTCGDLQGQDDDDLIYRGTKEWPSK 118  
 Qy 65 RK---NOYGSAAHVSLDDKD--DLVEKFSDIVNKAELPIRKONHIDLIDVFRKRGN 118  
 Db 61 SGPDLSKLMQASASDWMMENENCYNTSVPEUTGRRFLSPSLSPDMNLASAFKRRQ-- 118  
 Qy 119 RLHARTVQLQYETRKLUKOMENVIALSTTYTKETTCGDLQGQDDDDLIYRGTKEWPSK 178  
 Db 119 RLHARTVQLQYETRKLUKOMENVIALSTTYTKETTCGDLQGQDDDDLIYRGTKEWPSK 178  
 Qy 179 SNPYVNGDFTDRKGKGLLEVILLIYLAPNVAFLNRGHEDSTMARYGFIRESK 238  
 Db 179 ETPYVNGDFTDRKGKGLLEVILLIYLAPNVAFLNRGHEDSTMARYGFIRESK 238  
 Qy 239 YPKHGEREFLQFQDVISSLPVATIDGKILIVHGGSISQDUDFLRSLAELSHKVKHSEKLR 298  
 Db 239 YPKHGEREFLQFQDVISSLPVATIDGKILIVHGGSISQDUDFLRSLAELSHKVKHSEKLR 298  
 Qy 291 -----YVSLRPLL----- 299  
 Db 299 HRQIPINRKCRKQDSTASNSSSSSSSSSSISLGSLKTPCTHTPHRHPSSPTL 358  
 Qy 300 -----TIDGERPDKTCEVQIQFIMMSPDQATMGCVENTLRAAGYWGPDVT 344  
 Db 359 CGSGAQPQFLDLSASAPSPPHQEQWQKQVIDLWSPDQKAPNGCSPNTRGGCYFGPDVT 418  
 Qy 345 DNFLQRHLSVIRSHECKPQHFMHDWKITISASNYTAIGSKAYIRLNQMLPH 404  
 Db 419 RRLILOGQGLQQLRHSRCKQGQYELCHSGQVITISASNYTEEGSNRGAYIKVGRLMPR 478  
 Qy 405 FWQYISASQTRKLSPTRQMGTVESSALKLAVRURDRHDLBESFRKYDQGYSIS 464  
 Db 479 FVQY-QSFRKLTITQTRVAAKSALARKEKLFTHRSLILQFOQDHNNTGILV 537  
 Qy 465 HKCKVNUVNTKLGPLPHRDLRKLAPTSOKVNTTLDLTDVILEAEGMSMTDAL 524  
 Db 538 EWAQVLTGLBLPLWPKTRPHILASDGRVEYNCSFEDMEPGIPLVQTPNLA-BAL 595  
 Qy 525 YANKASLVAIFIIDANSGCTIDFETAIDLVAMHPGAVSKALEMKCRMMDNGD 584  
 Db 596 FRYKQKJEIIFNIKDHQGLSISERFRHTWHLFSAHLGKUDNRAIDLRSIDFNKDG 655  
 Qy 585 KUDNKFLEATRL 597  
 Db 656 NDNFNFELEATRV 668

RESULT 14  
 ID Q4RUL1\_TRTNG  
 AC Q4RUL1  
 DT 13-SEP-2005 (TREMBlre. 31, Created)  
 DT 13-SEP-2005 (TREMBlre. 31, Last sequence update)  
 DT 13-SEP-2005 (TREMBlre. 31, Last annotation update)  
 DB Chromosome 12 SCAT14996, whole genome shotgun sequence.

(Fragment).  
 ORIGIN=GSTENG0028998001;  
 OS Tetraodon nigroviridis; (Green puffer).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostei; Neoteleostei; Actinopterygii; Teleostei; Tetraodontiformes; Acanthomorpha; Tetraodontidae; Tetraodon.  
 OC Tetraodontoidea; Tetraodontidae; Tetraodon.  
 NCBI\_TaxID=99883;  
 RN [1]  
 RPP NUCLEOTIDE SEQUENCE.  
 RA Jallion O., Aury J.M., Brunet F., Petit J.L., Strange-Thomann N., Maucler E., Bouneau L., Fischer C., Ozouf-Costaz C., Bernot A., Niclaud S., Jaffe D., Fisher S., Dutfalla G., Bousat C., Segurens B., Basilia C., Salanoubat M., Levy M., Boudet N., Castellano S., Anthouard V., Jubin C., Castelli V., Kalinka M., Vacherie B., Blémont C., Skalli Z., Cattolico L., Poulin J., De Berardinis V., Cruaud C., Duprat S., Brottier P., Coutanceau J.P., Gouzy J., Parra G., Lardier G., Chapelle C., McEwan P., Bosak S., Kellis M., Volff J.N., Guigo R., Zody M.C., Mesirov J., Lindblad-Toh K., Birren B., Nusbaum C., Kahn D., Robinson-Rechavi M., Laude V., Schachter V., Quétier F., Saurin W., Scarpelli C., Wincker P., Lander E.S., Weissenbach J., Roest Crollius H., "Genome duplication in the teleost fish Tetraodon nigroviridis reveals the early vertebrate proto-karyotype.", Nature 431:946-957 (2004).  
 [2]  
 RPP NUCLEOTIDE SEQUENCE.  
 RG Genoscope; Whitehead Institute Centre for Genome Research;  
 RL Submitted (FEB-2004) to the EMBL/GenBank/DDBJ databases.  
 CC -> CARTION: The sequence shown here is derived from an EMBL/GenBank/DDBJ whole genome shotgun (WGS) entry which is preliminary data.  
 DR EMBL: CAAB0104986; CAGC08121.1; -; Genomic\_DNA.  
 FT NON\_TER 1  
 FT SEQENCE 797 AA; 797 MM; 89F01E058CB20730 CRC64;  
 Best Local Similarity 35.8%; Score 1232; DB 2; Length 797;  
 Matches 261; Conservative 135; Mismatches 190; Indels 218; Gaps 13;  
 Qy 5 AIRAAIFOKWYRQHARREMRQCNKQIFQNLKAYASRQDQAEYLKPRNDLKHMPQAG 64  
 Db 1 ALRAAVLFLQRWYRQYVARSBRRCRTWILFQSTIESGQAOIQTLISNFTLYLMDNFTPSN 60  
 Qy 65 RKNQYQGSAAHVSVLLDD-KDLDVEFGDINAKL-----ELPIRKNHIDLLIDVER 113  
 Db 61 ERNLI-- SHIFFENEVERGDAEHWERYFCKWIVPEYSGPHITPPLTVEQAVLVEAFR 117  
 Qy 114 KKGNGNLHAPKYYVALTRERAKSKJLKLPMNISPVTEAVSQQTV----- 156  
 Db 118 NKK--QLHSRYILQQLLSEIWKLMLPMLNRISTCHSKRITCGBLOGGNGLCRBI 175  
 Qy 157 -----GDLHGKDDLLAVLHK----- 172  
 Db 176 GFTLPLPKTVLRTWVAAPIGDHLQGLDELDLILIVYKVTRNPLPLPFLNSPPTVWLHVTYCPSN 235  
 Qy 173 -----NGLPPSSSNPYPVFNQDFVDRGKRGLEVLVLLLSLYLAFFNAFLR 217  
 Db 236 ANKKGKPNPVRNCALQNGMPSLREYVVFQDFVDRGDSIBELVILPARLIVVPSDVILR 295  
 Qy 218 GNFHDSSVMNAYGTPREVESKYPRNKHILAFDDEVYVPLPGSULSVRLVHGGSDS 277  
 Db 296 GNHEDHTVNLRYGTFKEVLYKVKHGRKILKLIQKIFSWPLATVIDOKVVLHGGISDL 355  
 Qy 278 TSUDLIKISRDKVSIIRPP----- 305  
 Db 356 TDIGLIAKURHNTYVSLRPKPKRGYNAGTSIDSVDVBDSVCNSKTFQRSLTIVARPL 415  
 Qy 306 D----- 307  
 Db 416 GTRTSFQNRLSQDFSDRTRIPAPAENELELRRRREPFIHLSESKEKTINPAASSPSVKSLSK 475

RESULT 15  
 QSMR2\_CAEEL  
 ID QMYR2\_CAEEL PRELIMINARY; PRT; 572 AA.  
 AC QMYR2;  
 DT 01-OCT-2002 (TREMBLrel. 22, Created)  
 DT 01-OCT-2002 (TREMBLrel. 22, Last sequence update)  
 DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)  
 DR Phosphatase with et hands protein 1, isoform b.  
 Name=pef-1; ORFlames=F23H11.8;  
 OS Caenorhabditis elegans;  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;  
 OC Rhabditidae; Peledorinae; Caenorhabditis.  
 NCBI\_TaxID=6239;

RN [1]  
 RPP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].  
 RC STRAIN=Bristol N2;  
 RX MEDLINE=9905613; PubMed=851916;  
 RG The C. elegans sequencing consortium;  
 RT "Genome sequence of the nematode *C. elegans*: a platform for investigating biology.", Science 282:12012-2018 (1998).  
 RL EMBL: AF003389; AACM22065.1; -; Genomic\_DNA.  
 DR HSSP; P36833; 1-16.  
 DR Ensembl; F23H11.8; Caenorhabditis elegans.  
 DR WormBase; WBGeno00001969; pef-1.  
 DR WormPep; F23H11.8b; CB30662.  
 DR GO; GO:0005509; F:calcium ion binding; IBA.  
 DR GO; GO:0016787; F:hydrolase activity; IEA.  
 DR InterPro; IPR01992; BP\_Hand\_Type.  
 DR InterPro; IPR00048; BP\_hand\_Ca\_bd.  
 DR WormPep; F23H11.8b; CB30662.  
 DR InterPro; IPR00843; M\_Pesterase.  
 DR InterPro; IPR00186; T\_phosphate\_apaH.  
 DR Pfam; PF00036; ethand.  
 DR Pfam; PF00149; Metallophospho; 1.  
 DR PRINTS; PRO014; S1PBPMSB.  
 DR SMART; SM00054; EFh; 3.  
 DR SMART; SM00156; PP2AC; 1.  
 DR PROSITE; PS00151; EF\_HAND; UNKNOWN\_2.  
 DR PROSITE; PS00125; SER\_THR\_PHOSPHATASE; UNKNOWN\_1.  
 KW Complete proteome.  
 SQ SEQUENCE 572 AA; 65150 MM; 504039A0685C98CD CRC54;  
 Query Match 33.3%; Score 1147; DB 2; Length 572;  
 Matches 227; Conservative 106; Mismatches 174; Indels 46; Gaps 10;  
 Qy 81 KODLVEERGFDINAKI-----LPIRKNHIDLLIDVFRKKGRHLR 121

Db	27 KPSLMDSGEGTVKKMLEDTSPVNDIDRNYKGFTLSPFLDKPQAVKMLAF - KVNVKLH	84	QY	3 ENIRAIATIQQWYRHOARZEMORRNQWQFQNLEYASSODAQSLYKPFNDLIC---- 5
QY	122 PKVVALILREBAAKSLKQLNISPVSTAVSQVTCGDLAHLGKQDLYLVTLHKGLPPSSNP	181	Db	14 EPLKALIIQWYRQKARJQVLTFRQSTEVADCGQMLSTPFSMPENTAHV 7
Db	85 PKVYMLILHEARTKFFKAMPSVRISTISNOVITCQGDLAGKEDOCLILYKONGYPVDNP	144	QY	58 HNPQAGRKNOYQGSATVSVLDKDVDEEVDIVAKTBLPIRKHIDLIIDVFRKKG 1
QY	182 YVFNGDFVDRKGKQLEUVLLSILYAPNAVLQRGKHEDSMARYGPIREVESKPR	241	Db	74 HKEGLELRNOSLSESEQDLRDRHDYVLDLDPDSYNGPRLQPLCTDIDLILBAFKEQQ- 1
Db	145 YIFNGDFVDRGGOSIIVLWLTVIPMSIVYANGRHEHDHMLRVGPIKLSKTYKD	204	QY	118 NELHPKVVALIRAEAKSLKQLNISPVSTAVSQVTCGDLAHLGKQDLYLVTLHKGLPS
QY	242 NHKRILAFIDDEVVWPLGSVLSRVLIVHGGSDSTSBDLKSDRKYVSLRPLTD	301	Db	133 -LHARHVLFVLFETKVKLKPQNFTHVOTSPSKELTICGDLAGKUDLFWIIFYKQGLPS
Db	205 LSPTRILIEDVPSWLPATIIDIIDPFDVHGIGISDQTEVKSLDKIPHRFOSVLRPVNK	264	QY	265 GMESKEKNSAVNVDWKOMLDIMWDSPKONKGCPWNVFRGGSYFFGADITASPLEKHGF
QY	302 GEPDK-----TEWOOFDIMWDSPDQATMGCVPTURGAGSYWGPDTDFLQRHRLS	354	Db	355 YVRSHECKPNGRPMHDNKKIITFSASNYIAQGNGKAYIR-LANOLMPHIVQYSAAS
Db	325 LLVRSHECKPEGFHSHNTCLTVFSSASYTGGSGRAGVYRPIGKSQPHVQYM-AS	382	QY	325 LLVRSHECKPEGFHSHNTCLTVFSSASYTGGSGRAGVYRPIGKSQPHVQYM-AS
QY	414 QTKRLS-PKORMGTVESSALKELAVMRDHDRDILDEFRKYDPKDGSYISTISHWCKYEN	472	Db	443 ITGLNLWPMLAPKAVTLEDGKVMYKDRRIAQVGCTHAQED--IYESLYRHKSTL
Db	383 KTHRKSTLBERLGVVEESAVKELKEKLGSSFPHTQDQKEFIMDTEKSGKLUPLKWSQYVER	442	QY	532 VAJFNIIADNSGATLDBFETADLVLVHMPGAYSKEM----L8KCRMMDLNNGKGV
QY	473 VTKLGLPMLLRDKLAP-GTDSQKVNMYRTLDLDTDYLTEAEDGMSVMDLYANCA	531	Db	500 ETLLPFDKMDNSQVSMSGPIDACEVIL---GKTYKPLQDTSQIAESIDFPNKQFPI
Db	443 ITGLNLWPMLAPKAVTLEDGKVMYKDRRIAQVGCTHAQED--IYESLYRHKSTL	499	QY	587 DLMEPFLEATRLSD 599
Db	555 DINSLLEAFRIVD 567		Db	329 PTLLRGAJGVWFGPDVTDNFQHRL
Db			QY	372 PTGRRGGCCYCPDPVTISKLMNQ
RESULT 16	QR441_MACFA	PRT: 405 AA.	RESULT 17	Q68EPO_XENTR
QR441_MACFA	ID: QR441_MACFA PRELIMINARY;	PRT: 405 AA.	ID: Q68EPO_XENTR PRELIMINARY;	PRT: 493 AA.
AC	QR441;		AC: Q68EPO,	
DT	13-SEP-2005 (TREMBlE, 31, Last sequence update)		DT: 068EPO,	
DT	13-SEP-2005 (TREMBlE, 31, Last annotation update)		DT: 068EPO,	
DE	Testis cDNA clone: Otsa_12861, similar to human protein phosphatase, DB hand calcium-binding domain 1 (PPB1), transcript variant 1.		DT: 25-OCT-2004 (TREMBlE, 28, Last sequence update)	
DB	Macaca fascicularis (crab eating macaque) (Cynomolgus monkey).		DT: 25-OCT-2004 (TREMBlE, 28, Last annotation update)	
OS	Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Buteleostomi; OC		DB: PPPSC-prov protein;	
OC	Mammalia; Rutheria; Rhachontoglires; Primates; Catarrhini; OC		DB: Name=ppp5C-prov;	
OC	Cercopithecidae; Cercopithecinae; Macaca.		DB: Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;	
OX	NCBI_TAXID=9541;		DB: OK	
RN	NUCLEOTIDE SEQUENCE.		RN: Nucleotide Sequence.	
RP	NUCLEOTIDE SEQUENCE.		RP: Nucleotide Sequence.	
RC	TISSUE=Embryo;		RC: Tissue=Embryo;	
RX	MEDLINE=22388257, PubMed=12477932, DOI=10.1073/pnas.242603899;		RX: MEDLINE=22388257, PubMed=12477932, DOI=10.1073/pnas.242603899;	
RA	Strauberg R.L., Peingold B.A., Grouse L.H., Derge J.G.,		RA: Strauberg R.L., Peingold B.A., Grouse L.H., Derge J.G.,	
RA	Klaushuber R.D., Collins F.S., Wagner L., Sheinman C.M., Schuler G.D.,		RA: Klaushuber R.D., Collins F.S., Wagner L., Sheinman C.M., Schuler G.D.,	
RA	Altschul S.F., Zeeberg B., Buttow K.H., Schaefer C.F., Bhat N.K.,		RA: Altschul S.F., Zeeberg B., Buttow K.H., Schaefer C.F., Bhat N.K.,	
RA	Hopkins R.P., Jordan H.H., Moore T., Max S.I., Wang J., Heilek P.,		RA: Hopkins R.P., Jordan H.H., Moore T., Max S.I., Wang J., Heilek P.,	
RA	Blatchko L., Maruska K., Farmer A.A., Rubin G.M., Hong L.,		RA: Blatchko L., Maruska K., Farmer A.A., Rubin G.M., Hong L.,	
RA	Stapleton M., Soares M.B., Ronald M.F., Cassavant T.L., Scheetz T.E.,		RA: Stapleton M., Soares M.B., Ronald M.F., Cassavant T.L., Scheetz T.E.,	
RA	Brownstein M.J., Usdin T.B., Yoshiaki S., Carnici P., Range C.,		RA: Brownstein M.J., Usdin T.B., Yoshiaki S., Carnici P., Range C.,	
RA	Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullay S.J.,		RA: Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullay S.J.,	
RA	Bosak S.A., McElvan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,		RA: Bosak S.A., McElvan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,	
RA	Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,		RA: Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,	
RA	Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,		RA: Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,	
RA	Fahy J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,		RA: Fahy J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,	
RA	Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,		RA: Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,	
RA	Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,		RA: Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,	
RA	Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,		RA: Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,	
RA	Sugano S., Gojobori T., Shen J.C.-K., Wu C.I., Hashimoto K.,		RA: Sugano S., Gojobori T., Shen J.C.-K., Wu C.I., Hashimoto K.,	
RA	"DNA sequences of macaque genes expressed in brain or testis and its evolutionary implications.",		RA: "DNA sequences of macaque genes expressed in brain or testis and its evolutionary implications.",	
RA	Submitted (JUN-2005) to the EMBL/GenBank/DDBJ databases.		RA: Submitted (JUN-2005) to the EMBL/GenBank/DDBJ databases.	
RA	[2]		RA: [2]	
RA	NUCLEOTIDE SEQUENCE.		RA: Nucleotide Sequence.	
RA	International consortium for macaque cDNA sequencing, analysis, and its evolutionary implications.",		RA: International consortium for macaque cDNA sequencing, analysis, and its evolutionary implications.",	
RA	Submitted (JUN-2005) to the EMBL/GenBank/DDBJ databases.		RA: Submitted (JUN-2005) to the EMBL/GenBank/DDBJ databases.	
RA	[2]		RA: [2]	
RA	Best Local Similarity 41.3%; Pred. No. 1e-42; Length 405;		RA: Best Local Similarity 41.3%; Pred. No. 1e-42; Length 405;	
RA	March 159: Conservative 79; Mismatchs 111; Index 35; Gap 4.		RA: March 159: Conservative 79; Mismatchs 111; Index 35; Gap 4.	
RA	Submittted (MAY-2004) to the EMBL/GenBank/DDBJ databases.		RA: Submittted (MAY-2004) to the EMBL/GenBank/DDBJ databases.	
RA	EMBL: AB17904; BAE02125.1; -; MIGRA.		RA: EMBL: AB17904; BAE02125.1; -; MIGRA.	
RA	SEQUENCE 405 AA; 46636 MN; CIE9889F2023E4B CRC64;		RA: SEQUENCE 405 AA; 46636 MN; CIE9889F2023E4B CRC64;	

CC phosphate. Belongs to the PPP phosphatase family.  
 CC -!- SIMILARITY: EMBL; BC00162; AAB80162.1; - ; mRNA.  
 DR EMBL; Q65BP0; 20-433.  
 DR Ensemble; ENSEMBL; G0005737; Cyttoplasm; IEA.  
 DR GO; GO:0005737; Cytoplasm; IEA.  
 DR GO; GO:0005634; C-nucleus; IEA.  
 DR GO; GO:0016787; Hydrolase activity; IEA.  
 DR GO; GO:0006471; P-Phosphoprotein phosphatase activity; IEA.  
 DR InterPro; IPR004943; M-peptidase.  
 DR InterPro; IPR01236; Pptase\_5'.  
 DR InterPro; IPR006186; T\_phthase\_apah.  
 DR InterPro; IPR01440; TPR.  
 DR InterPro; IPR01990; TPR-like\_helical.  
 DR Pfam; PRO0149; Metallophos; 1.  
 DR Pfam; PRO0515; TPR\_1; 3.  
 DR PROSITE; PS00005; TPR\_3.  
 DR PROSITE; PS00293; TPR\_REGION; 1.  
 KW Hydrolase; Iron; Repeat; TPR repeat.  
 SQ SEQUENCE 493 AA; 56247 MW; DD7597E849DD40ER CRC64;

Query Match 18.0%; Score 618.5%; DB 2; Length 897;  
 Best Local Similarity 31.2%; Pred. No. 7.2e-25; Matches 190; Conservative 88; Mismatches 215; Indels 115; Gaps 19;  
 Matches 190; Pred. No. 7.2e-25; Mismatches 215; Indels 115; Gaps 19;

QY 92 VVAKIEPPIRKNDLIDVFRKGRHLPPKVAL--ILRRAAKSKQLMNPSTAV 149  
 DR 182 IDFWELBLMFGYDQOKKLHCKCVCYQMLVWYDKIISKLPSLVEISLEKSQQTVCGPTGQP 241  
 DR 164 DDLILVHLHKNGLPSSSNPVFNGDFDVDRGRLGVLLLSSLYLAFFNPAVFLNRGNHEDS 223  
 DR 242 YDLANIPIHNGLSENNPFYFNGDFDVDRGSPSVEVIVILRGPKLILYPAQPHLQRNGHTD 301  
 DR 224 VMMARYGFTREVSKYPRHKRILAFPIKVYMLPLGSVTLNSVNLIVIGGP--SDSTSLD 281  
 DR 302 TMNQMYGFGEGBVKARY--SAQMQLFSLFVQMLPLAMCVNQRLVIMIGGLFSBDGVTD 358  
 DR 282 LIKSIDRKGKXVISLRRPDTGEPBLDKTWWQIQIDIMWMSDPOATMGCVPTNLQAGVWFGP 341  
 DR 359 QINSIRURR----QPP----DSGPK-----CDLWSDPDPQDG-KSSSKKGIVSGCPGP 402  
 DR 342 DVTDNFLQHRLSVIRSHCKNGHETPHMDNKITITASANNTAIGSNKGAYTRIN-NQ 400  
 DR 403 DVTRRFLENGLGLAYTIRSHENKQEGYEVSHINGLCVTVISAPNYCDQMGNKGAYTHNGSD 462  
 QY 401 LMPHPVQY 408  
 DR 463 LKPKFHQF 470

RESULT 18  
 Q0QGM7 LEIMA PRELIMINARY; PRT; 897 AA.  
 AC Q0QGM7; LEIMA PRELIMINARY;  
 DT 13-SEP-2005 (Tremblrel. 31, Created)  
 DT 13-SEP-2005 (Tremblrel. 31, Last sequence update)  
 DT 13-SEP-2005 (Tremblrel. 31, Last annotation update)  
 DB Serine/threonine protein phosphatase-like protein (BC 3.1.3.16).  
 GN ProName=ImpP12\_0660;  
 OS Leishmania major.  
 OC Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.  
 NCBI\_TaxId=564;  
 RN [1] NUCLEOTIDE SEQUENCE.  
 RP STRAIN=tryedin;  
 RA Peacock C.S., Murphy L., Ivens A.C., Berriman M., Blackwell J.,

RA Smith D., Collins M., Foster N., Harris D., Oliver K., O'Neil S.';  
 RA Saunders D., Seeger K., Warren T., Rajandream M., and Barrell B.G.;  
 RL Submitted (JUN-2005) to the EMBL/GenBank/DDBJ databases.  
 DR EMBL; Cn00251; CAJ02710.1; - ; Genomic\_DNA.  
 KW Hydrolase.

SEQUENCE 897 AA; 10352 MW; 1631BB67C70DAF1 CRC64;

Query Match 18.0%; Score 618.5%; DB 2; Length 897;  
 Best Local Similarity 31.2%; Pred. No. 7.2e-25; Matches 190; Conservative 88; Mismatches 215; Indels 115; Gaps 19;  
 Matches 190; Pred. No. 7.2e-25; Mismatches 215; Indels 115; Gaps 19;

QY 92 VVAKIEPPIRKNDLIDVFRKGRHLPPKVAL--ILRRAAKSKQLMNPSTAV 149  
 DR 182 IDFWELBLMFGYDQOKKLHCKCVCYQMLVWYDKIISKLPSLVEISLEKSQQTVCGPTGQP 241  
 DR 164 DDLILVHLHKNGLPSSSNPVFNGDFDVDRGRLGVLLLSSLYLAFFNPAVFLNRGNHEDS 223  
 DR 242 YDLANIPIHNGLSENNPFYFNGDFDVDRGSPSVEVIVILRGPKLILYPAQPHLQRNGHTD 301  
 DR 224 VMMARYGFTREVSKYPRHKRILAFPIKVYMLPLGSVTLNSVNLIVIGGP--SDSTSLD 281  
 DR 302 TMNQMYGFGEGBVKARY--SAQMQLFSLFVQMLPLAMCVNQRLVIMIGGLFSBDGVTD 358  
 DR 282 LIKSIDRKGKXVISLRRPDTGEPBLDKTWWQIQIDIMWMSDPOATMGCVPTNLQAGVWFGP 341  
 DR 359 QINSIRURR----QPP----DSGPK-----CDLWSDPDPQDG-KSSSKKGIVSGCPGP 402  
 DR 342 DVTDNFLQHRLSVIRSHCKNGHETPHMDNKITITASANNTAIGSNKGAYTRIN-NQ 400  
 DR 403 DVTRRFLENGLGLAYTIRSHENKQEGYEVSHINGLCVTVISAPNYCDQMGNKGAYTHNGSD 462  
 QY 401 LMPHPVQY 408  
 DR 463 LKPKFHQF 470

RESULT 19  
 Q042205 XENLA  
 ID 042205\_XENLA PRELIMINARY; PRT; 492 AA.  
 AC Q042205; XENLA PRELIMINARY;  
 DT 01-JAN-1998 (Tremblrel. 05, Created)  
 DT 01-JAN-1998 (Tremblrel. 05, Last sequence update)  
 DT 01-MAR-2004 (Tremblrel. 26, Last annotation update)  
 DB Protein phosphatase 5 (Fragment).  
 Name=pp5;  
 OS Xenopus laevis (African clawed frog).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Buteleostomi;  
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;  
 OC Xenopodinae; Xenopus; Xenopus.  
 NCBI\_TaxId=8355;  
 RN [1] NUCLEOTIDE SEQUENCE.  
 RP Ollendorff V., Donoghue D.J.;

RL Submitted (AUG-1997) to the EMBL/GenBank/DDBJ databases.  
 CC -|- CATALYTIC ACTIVITY: A phosphoprotein + H(2)O = a protein +  
 phosphate.  
 CC phosphate.  
 CC -|- SIMILARITY: Belongs to the PPP phosphatase family.  
 CC  
 DR EMBL: AF018263; ARB:70574.1; -; mRNA.  
 DR P53041; LA17.  
 DR SNR; 042205; 19492.  
 DR GO; GO:000537; C:cytoplasm; IEA.  
 DR GO; GO:0005634; C:nucleus; IEA.  
 DR GO; GO:0005688; F:binding; IEA.  
 DR GO; GO:001687; F:hydrolyase activity; IEA.  
 DR GO; GO:0004721; F:phosphoprotein phosphatase activity; IEA.  
 DR GO; GO:006570; P:protein amino acid dephosphorylation; IEA.  
 DR InterPro; IPR001843; M-peptidase.  
 DR InterPro; IPR011236; PPtase\_5.  
 DR InterPro; IPR006186; T\_phosphatase\_apAH.  
 DR InterPro; IPR001440; TPR.  
 DR InterPro; IPR011990; TPR-like\_helical.  
 DR Pfam; PF00149; Metallophos; 1.  
 DR Pfam; PF00515; TPR; 1; 3.  
 DR PROSITE; PS5005; TPR; 3.  
 DR PROSITE; PS5029; TPR\_REGION; 1.  
 DR PRODOM; PD000252; T\_phosphatase\_apAH; 1.  
 DR SMART; SMD0156; PP2AC; 1.  
 DR SMART; SMD0028; TPR; 3.  
 DR PROSITE; PS50125; SER\_THR\_PHOSPHATASE; UNKNOWN\_1.  
 DR PROSITE; PS5005; TPR; 3.  
 DR PROSITE; PS5029; TPR\_REGION; 1.  
 KW Hydroxase; Iron; TPR repeat.  
 FT NON\_TER 1 1  
 SQ SEQUENCE 492 AA; 56145 MW; 868P0641C33DD16 CRC64;  
 Query Match 17.9%; Score 615.5; DB 2; Length 492;  
 Best Local Similarity 42.9%; Pred. No. 5e-29; Mismatches 103; Indels 23; Gaps 8;  
 Matches 132; Conservative 50; Mismatches 103; Indels 23; Gaps 8;  
 Qy 105 IDLLIDVFR-KKGRNRLHPKTVALLIRRAAKSLQQLNPVSTAVSQQTVCQGDHLGKL 163  
 Db 181 VDFMLBLMOPYKQDKKLCKLWOMVQYDQTLKSLPSLVTSLSRSQQVTCGDTGQP 240  
 Qy 164 DDLILVVLKNGLQSSNSNYPVENDFVDRGKGRLGEVLVLLSLYLAFNAHVNLRGNHEDS 223  
 Db 241 YDLMNIPHLNGLISRNPNPYIFNDFVDRGGSFSVEVITLGKFLVLAFFHLRGHETD 300  
 Qy 224 VMAPYSPFIREYESVKPRHNKRILAFIDEVYVWPLQGSVLSVNLVLTHGGR--SDSTSLD 281  
 Db 301 TWMQMYGPGEPEVKAYK--SAQMFOLFSEVRQWQPLAMCNVQRVLMHGIGPSEDGVLD 357  
 Qy 282 LTKSIDGCKYKVSLRPPTDGERPLKTEWQOFTDIMSPPATMGCVPTNTRGAGWFGP 341  
 Db 358 QIKRNIEKR----OPP--DGGM-----CDLWMSDPDOPG-RSTSRRGVSCQGP 401  
 Qy 342 DVTDNFLORHKSYVIRSHBCKENGHFRPHMDNKLITFSASVYALGSNKAYIRAN-NQ 400  
 Db 402 DVTQHOFLELNGDYTIRSHEVKEPEGEVSHGLCVTVFSAVNYCDQMGKAYIHLGSD 461  
 Qy 401 LMPHFMVOY 408  
 Db 462 LKPKFKHOP 469  
 RESULT 20  
 O6GPS6 XENLA  
 ID O6GPS6\_XENLA PRELIMINARY; PRT; 493 AA.  
 AC 06GPS6;  
 DT 05-JUL-2004 (TREMBlrel. 27, Last sequence update)  
 DT 05-JUL-2004 (TREMBlrel. 27, Last annotation update)  
 DE Pp5\_protein.  
 GN Name=pp5;  
 OS Xenopus laevis (African clawed frog).  
 OC Bokaryota; Metazoa; Chordata; Craniota; Vertebrata; Ruteleostomi;  
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;

OC Xenopodine; Xeropus; Xencopus.  
 NCBI\_TaxID=8355;  
 [1]  
 RN NUCLEOTIDE\_SEQUENCE.  
 RC TISSUE=Ovary;  
 RX MEDLINE=22388217; PubMed=12477932; DOI=10.1073/pnas.242603999;  
 RA Straubberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klaunser R.D., Collins P.S., Wagner L., Shearer C.M., Schuler G.D.,  
 RA Altshull S.F., Zeeberg B., Bustow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.P., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Ronald M.P., Casavant T.L., Scheetz T.B.,  
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Raha S.S., Loquaiello N.A., Peters G.J., Abramson R.D., Mullahy S.J.,  
 RA Bosak S.A., McElroy P.J., McKernan K.J., Malek J.J., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Rulyk S.W.,  
 RA Villaon D.K., Murzny D.M., Sodergren B.J., Lu X., Gibbs R.A.,  
 RA Fahy J., Helton E., Kattelman M., Madan A., Rodriguez S., Sanchez A.,  
 RA Whiting M., Madan A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green B.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smialius D.E.,  
 RA Schniech A., Schein J.B., Jones S.J.M., Marra M.A.,  
 RT "Generation and initial analysis of more than 15,000 full-length human  
 RT and mouse cDNA sequences";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).  
 RN [2]  
 RP NUCLEOTIDE\_SEQUENCE.  
 RC TISSUE=Ovary;  
 RX MEDLINE=2234113; PubMed=12454917; DOI=10.1002/dvdy.10174;  
 RA Klein S.L., Straubberg R.L., Wagner L., Pontius J., Clifton S.W.,  
 RA Richardson P.;  
 RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus  
 RT initiative";  
 RL Dev. Dyn. 225:384-391 (2002).  
 RN [3]  
 RP NUCLEOTIDE\_SEQUENCE.  
 RC TISSUE=Ovary;  
 RA Klein S., Straubberg R.;  
 CC Submitted (JUN-2004) to the EMBL/GenBank/DDBJ databases.  
 CC -|- CATALYTIC ACTIVITY: A phosphoprotein + H(2)O = a protein +  
 CC phosphate.  
 CC -|- SIMILARITY: Belongs to the PPP phosphatase family.  
 CC EMBL: BC073033; AAH73033.1; -; mRNA.  
 DR SMR; O6GPS6; 20-49.  
 DR GO; GO:0005377; C:cytoplasm; IEA.  
 DR GO; GO:0005634; C:nucleus; IEA.  
 DR GO; GO:001687; F:hydrolyase activity; IEA.  
 DR GO; GO:0004721; F:phosphoprotein phosphatase activity; IEA.  
 DR GO; GO:0006470; P:protein amino acid dephosphorylation; IEA.  
 DR InterPro; IPR008443; M-peptidase.  
 DR InterPro; IPR011236; PPtase\_5.  
 DR InterPro; IPR006186; T\_phosphatase\_apAH.  
 DR InterPro; IPR001440; TPR.  
 DR InterPro; IPR011990; TPR-like\_helical.  
 DR Pfam; PF00149; Metallophos; 1.  
 DR PROSITE; PS5005; TPR; 3.  
 DR PROSITE; PS5029; TPR\_REGION; 1.  
 DR Hydrolase; Iron; Repeat TPR repeat.  
 DR SMART; SMD0156; PP2AC; 1.  
 DR SMART; SMD0028; TPR; 3.  
 DR PROSITE; PS50125; SER\_THR\_PHOSPHATASE; UNKNOWN\_1.  
 DR PROSITE; PS50005; TPR; 3.  
 DR PROSITE; PS5029; TPR\_REGION; 1.  
 DR Hydrolase; Iron; Repeat TPR repeat.  
 DR SMART; SMD0052; T\_phosphatase\_apAH; 1.  
 DR SMART; SMD0156; PP2AC; 1.  
 DR SMART; SMD0028; TPR; 3.  
 DR PROSITE; PS50125; SER\_THR\_PHOSPHATASE; UNKNOWN\_1.  
 DR PROSITE; PS50005; TPR; 3.  
 DR PROSITE; PS5029; TPR\_REGION; 1.  
 DR Hydrolase; Iron; Repeat TPR repeat.  
 SQ SEQUENCE 493 AA; 56276 MW; 127DC9409402C297 CRC64;  
 Query Match 17.9%; Score 615.5; DB 2; Length 493;  
 Best Local Similarity 42.9%; Pred. No. 5e-29; Mismatches 103; Indels 23; Gaps 8;  
 Matches 132; Conservative 50; Mismatches 103; Indels 23; Gaps 8;  
 Qy 105 IDLLIDVFR-KKGRNRLHPKTVALLIRRAAKSLQQLNPVSTAVSQQTVCQGDHLGKL 163

DR PROSITE; PS50293; TPR\_REGION; 1.  
 DR KW Hydrolase; Iron; TPR\_Repeat.  
 DR FT 1  
 DR NON\_TER; 1  
 DR SEQID 474 AA; 54287 MW; A7BDD4FD5BCAC17A CRC64;  
 DR  
 DR Query Match 17.9%; Score 614.5; DB 2; Length 474;  
 DR Best Local Similarity 36.4%; Pred. No. 5\_se-29; Mismatches 155; Indels 43; Gaps 12;  
 DR Matches 150; Conservative 64; Mismatches 155; Indels 43; Gaps 12;  
 DR SEQUENCE 474 AA:  
 DR 11 PIQKWRHQAERREMRQRNCWQFONIYAYA-SQDQAELYKFENDLKIKNPQAAGR 65  
 DR 74 YTKAVYRRAALMALGR--FKKALADIEFVAKRCPSAKAQDKYTECKKVNKLAFSEKAI 131  
 DR 66 KNOYQGSAHSVUVDP---KDLVSEFGDIYNAKELPPIRKHILLIDVPRKGRR 119  
 DR 342 DVTDNFLQRHLSVIRSHCECKNGHEFMHDNKITISASNTAIGSNKGAYIRLN\_QO 400  
 DR 403 DVTHQFLENGGLYVIRSHCEVKRGYEVSHNLGVTCVTFSAFPNCQDMNKNGATIHSOD 462  
 DR 401 LMPHIVQY 408  
 DR 463 LKPKEFHQP 470

RESULT 21

DR Q7OB71\_ANOGA

ID Q7OB71\_ANOGA PRELIMINARY; PRT; 474 AA.

AC [2]

DT 01-MAR-2004 (Tremblrel. 26, Created)  
 DT 01-MAR-2004 (Tremblrel. 26, Last sequence update)

DE ENSANGP0000011234 (Fragment).

GN Anopheles\_gambiae SRR. PEST.

OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea; Culicidae;  
 OC Anophelinea; Anopheles.  
 NCBI\_TAXID=180454;

RN [1]

RP NUCLOBOTIDE SEQUENCE.

RC STRAIN-PEST:

RT The Anopheles gambiae Sequence Committee,  
 "Anopheles gambiae re-annotation.",  
 Submitted (APR-2002) to the EMBL/GenBank/DDBJ databases.

RL [2]

RN [1]

RP NUCLOBOTIDE SEQUENCE.

RC STRAIN-PEST:

RT The Anopheles gambiae Sequence Committee,  
 Submitted (APR-2004) to the EMBL/GenBank/DDBJ databases.

-1 CATALYTIC ACTIVITY: A phosphoprotein + H<sub>2</sub>O = a protein + phosphate.

-1 CAUTION: The sequence shown here is derived from an EMBL/GenBank/DDBJ whole genome shotgun (WGS) entry which is preliminary data.

-1 SIMILARITY: Belongs to the PPP phosphatase family.

EMBL; AAB01008890; EA08659; 2; -; Genomic\_DNA.

DR GO\_0005737; C:cytoplasm; IBA.

DR GO\_0005634; C:nucleus; IBA.

DR GO\_0016787; F:hydrolase activity; IBA.

DR GO\_0004721; F:Protein phosphatase activity; IBA.

DR GO\_0006470; F:Protein amino acid dephosphorylation; IBA.

DR InterPro; IPR00443; M-pesterase.

DR InterPro; IPR01236; ppptase\_5.

DR InterPro; IPR006186; T\_pttase\_apAH.

DR InterPro; IPR00140; TPR.

DR PRO0149; Metallophos; 1.

PFam; PF00515; TPR\_1; 2.

PRSR; PRSR033096; ppptase\_5; 1.

DR PR0014; STPPPTASE; 5; 1.

PR000252; T\_pttase\_apAH; 1.

SMART; SW00156; PP2Ac; 1.

DR SMART; SW0028; TPR; 2.

PROSITE; PS50005; TPR; 2.

RN [1]

RESULT 22

DR Q8RFT2\_TETNG

ID Q8RFT2\_TETNG PRELIMINARY; PRT; 874 AA.

AC [2]

DT 13-SEP-2005 (Tremblrel. 31, Created)  
 DT 13-SEP-2005 (Tremblrel. 31, Last sequence update)

DB Chromosome 16 SCAP15113, whole genome shotgun sequence.

DB (Fragment).

GN ORFNName=GSTENG0035185001,

OS Tetradon nigroviridis (Green puffer).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Buteleostomi; Actinopterygii; Neopterygii; Teleostei; Buteleosteii; Nectoleosteii; Acanthomorpha; Acanthopterygii; Percromorpha; Tetraodontiformes; Tetraodontidae; Tetraodontidae; Tetraodon.

OC Tetraodontidae; Tetraodontidae; Tetraodon.

OX NCBI\_TAXID=9883;

RN [1]

RP NUCLOBOTIDE SEQUENCE.

RA Jaijon O., Aury J.M., Brunet F., Petit J.-L., Stange-Thomann N., Mauceli E., Bouneau L., Blacher C., Ozouf-Cossat C., Berrot A., Nicaud S., Jaffe D., Fisher S., Lutella G., Brossat C., Seguens B., Basilia C., Salanoubat M., Levy M., Boudet N., Castellano S., Anthouard V., Jubin C., Castelli V., Katinka M., Vacherie B., Bleumont C., Skalli Z., Cattolico L., Poulain J., De Bernardini V., Cruaud C., Duprat S., Broto-P., Coutanceau J.-P., Gouzy J., Parra G., Hardier G., Chapelle C., McKernan K.J., McElwain P., Bosak S., Kellis M., Wolfson J.N., Guigo R., Zody M.C., Meirov J., Lindblad-Toh K., Birren B., Nusbaum C., Kahn D., Robinson-Rechavi M., Laudet V., Schachter V., Quétier F., Saurin W., Scarpelli C., Wincker P., Lander E.S., Weissenbach J., Roest Crollius H.;

RT "Genome duplication in the teleost fish Tetraodon nigroviridis reveals the early vertebrate proto-karyotype.";

RT Nature 431:946-957(2004).

[2]

RP NUCLEOTIDE SEQUENCE: Institute Centre for Genome Research;  
 RG Genoscope; Whitehead Institute Centre for Genome Research;  
 RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.  
 CC -I- CAUTION: The sequence shown here is derived from an  
 EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is  
 CC preliminary data.  
 CC -I- CATALYTIC ACTIVITY: A phosphoprotein + H(2)O = a protein +  
 phosphate.  
 DR InterPro; IPR004843; M-peptidase.  
 DR InterPro; IPR006186; T\_Ptase\_apAH.  
 DR InterPro; IPR004401; TPR.  
 DR InterPro; IPR006122; Znf\_C2CH.  
 DR PFAM; PF00149; Metallophob\_1.  
 DR PFAM; PF05405; TRAP\_1.  
 DR PFAM; PF00515; TPR\_1\_3.  
 DR PRINTS; PS00125; STPHPTASE.  
 DR PRODOM; P000252; T\_phptase\_apAH\_1.  
 DR SMART; SM00692; DM3\_1.  
 DR SMART; SM00156; PP2AC\_1.  
 DR SMART; SM00228; TPR\_3.  
 DR PROSITE; PS00125; SER THR PHOSPHATASE; UNKNOWN\_1.  
 DR PROSITE; PS0005; THAF\_1.  
 DR PROSITE; PS00293; TPR\_REGION\_1.  
 DR KW Hydrolase; Iron; TPR repeat.  
 FT NON\_TER 1  
 SQ SEQUENCE 874 AA; 98819 MW; E060861862DD0BFA CRC64;  
 Query Match 17.8%; Score 613; DB 2; Length 874;  
 Best Local Similarity 31.3%; Pred. No. 1.5e-28;  
 Matches 161; Conservative: 95; Mismatches 166; Indels 92; Gaps 21;  
 Qy 1 MDENAIARAAFIQKQWYRHQARREMR-RCNNWOLFQMLEYA--SEODQAELEYKFENDLIK 57  
 Db 53 IDRN-----YIKQYIRRAATSNMALKFKAALKDQEYTVRVRVPNDKQARMKIQQBCNCKV 106  
 Qy 58 HMMQAGRKKNQYQGSARHVSVDKDD---DLVERFG--DIVAKIELPIKRKHIDLIDV 111  
 Db 107 ---QKAFAERAIASDETKEKVSVVDSDIENNTIEDYAGPKLENKGKVLAFMKE---MMEW 159  
 Qy 112 FRKKRGNLHLPKVALILRAAALKSILQKLPNISPVSTAVSQVTVCGDGLRKGKDDLLVVLH 171  
 Db 160 FPKQK---KLHKCAVQIIVQVKOVLKSLPSLRLTKEVKKTICGTHQYDQYDUNIFK 217  
 Qy 172 KNGLSSSNPQYENGDFDRGRKGKLEVLILLISLYLAPPNAVFLNRGNHEDVMNARYGF 231  
 Db 218 LNSGPSETNPYLNGDFVDRGFSLEVLITLFGFKLUYPDNFHLLRGNHETDMNQYGF 277  
 Qy 232 IREVSQYPRNHRILALPIDEVTRMLPGSVMSRVLVHGP--SOSTSDLKSIDRQ 289  
 Db 278 EGEVKAKT---AQMFOLFSEVQWMLPLAQCNINKVLMHGGFLPSDQVTLSDRDKDRN 334  
 Qy 290 KVYSLIRPLDQBLDKTBWQQIFPDIMMSDPAQTMGCUPNTLRGAGVWGPDTDNPLQ 349  
 Db 335 R-----OPP---DSGPW-----CDLWSPDPQPONG-RSISRGVSCQPGDVTBPLS 378  
 Qy 350 RHLRSYVTRSHRECKPNHGFHMHDNKKITIFFASNYVATGSNKGAYVRL-NNGLOMPHVOY 408  
 Db 379 QNLQDFIVSRHETKAGEGEVTRSGKCTIVFSAFANYCDOMGNGKAYIHLRGSDQKPERHQ 438  
 Qy 409 ISASASOTKRLSKFQRMGVIVESSALKEA-----VNMRDHRDELDIBERKTDKPDQSY 460  
 Db 439 TAVASTTRH-----DLTTVHERAFTLFLQLVYQGDHRQOM-----PK--H 477  
 Qy 451 ISTSHWKYKMENTVIGLIPWMLRDKLAPGTDQ 494  
 Db 478 CSVPN-CK-----NGSGPQTDK 494

RP 04Y018 PLACH PRELIMINARY; PRT; 953 AA.  
 AC 04Y018; ID ID  
 DT 13-SEP-2005 (TREMBIrel. 31, Created)  
 DT 13-SEP-2005 (TREMBIrel. 31, Last sequence update)  
 DT 13-SEP-2005 (TREMBIrel. 31, Last annotation update)  
 DR PPI-1-like protein serine/threonine phosphatase, putative.  
 DR ORFNames=PC000121.02.0;  
 GN OS Plasmodium chabaudi.  
 OS Baktyota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.  
 OC NCBI\_TAXID=5825;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE.  
 RA Hall N., Karras M., Raine J.D., Carlton J.M., Koçiј T.W.A.,  
 RA Berriman M., Florens L., Janssen C.S., Pain A., Christopherides G.K.,  
 RA James K., Rutherford K., Harris B., Harris D., Churcher C.,  
 RA Quail M.A., Ormond D., Doggett J., Truman H.E., Mendoza J.,  
 RA Bidwell S.L., Rajandream M.A., Carucci D.J., Yates J.R., Karatasos F.C.,  
 RA Janse C.J., Barrell B., Turner C.M.R.P., Waters R.P., Sinden R.S.,  
 RA "A comprehensive survey of the Plasmodium life cycle by genomic,  
 RT transcriptomic, and proteomic analyses.",  
 RC Science 307:82-86 (2005).  
 CC -I- CAUTION: The sequence shown here is derived from an  
 CC preliminary data.  
 DR EMBL; CA0010923; CAH77292.1; -; Genomic DNA.  
 SQ SEQUENCE 953 AA; 111137 MW; 2C224CB0D1667B09 CRC64;  
 Query Match 17.5%; Score 603.5; DB 2; Length 953;  
 Best Local Similarity 26.9%; Pred. No. 6.3e-28;  
 Matches 196; Conservative: 122; Mismatches 264; Indels 147; Gaps 23;  
 Qy 8 AAIFIQKQWYRHQARREMR-CN WQIQLNL-EVASQDQABLY------ 49  
 Db 83 ACILQKIQYRGHARKFHEFTINCKWKRKFEDLHENIVLNHEIYVPLMKKINEDLONG 142  
 Qy 50 -----KEFDLKKMROAAGRKNQYQGSFHSVLDKDD---LVEFGDIVNAK-LELP 99  
 Db 143 IIQFNRYKPSNSNSRORHMROR--CN WQIQLNL-EVASQDQABLY------ 49  
 Qy 100 IRKHHDLLIDVFRKKGKGNRHLPKVALILRAAALKSILQKLPNISPVSTAVSQVTVCGD 159  
 Db 203 I-SMVKILNATKOMLSKNTKSVYLNLDKSKETK-----LVLGDV 245  
 Qy 160 HKGLDILVWHLKNGKPSNSPQYENGDFDRGRKGKLEVLILLISLYLAPPNAVFLNRGN 219  
 Db 246 HQQNDVWLWLFNRFGLSFSSTVYIENGDIADRGQNAABIFLWFAPKLSNVDVLRGN 305  
 Qy 220 HDSDVVAARYVIREVSKYRHLILALPIDEVTRMLPGSVMSRVLVHGGSDSTS 279  
 Db 306 HHCYSMYNEVYQFHNEVTSKVDSS--VFDIPOBIFELISLVSNTQIQIVFHGGSLRYQD 362  
 Qy 280 LDL-TKSIDRKYVSLIRPLDQBLDKTBWQQIFPDIMMSDPAQTMGCUPNTLRGAGV 337  
 Db 363 LSKVQDNLKCH\_EILHP-----EKEDTIPFLWSPDKQDGKQGQGNGNCI 413  
 Qy 338 WFGPDVMTDNFLQRHRLSSYVIRSHHECKP--NCFHEPMIDNKKITIFFASNYVATGSNKGAYI 395  
 Db 414 AFGPDITESFLKNNKDFFLIRSHQFVKLQIESHNGKCTIFSASNYCNKIKLGA 473  
 Qy 396 RLLNQIQLPHVQYIASASQTRLSF---KORMGIVESSALKE----- 434  
 Db 474 IFNQDITFPEVHRYMSPSLDVTFRETEFENOKURKVNNSNLEVEKBNRKAStPPDA 533  
 Qy 435 -----LAVMRDHRDBLEBEFRKIDPKDQSYISISHWCKMVENTKL\_GLPWLRDK 486  
 Db 534 LNDLNLFLSTLUCINEKLNMLNLYEKGKGIGEVINWTRBLKTKAKNVWPLICK 593  
 Qy 487 LAGTDSOKVNTNTLDDTDV---TIBABADG--MSVMDALYANKASLVALNTI 538  
 Db 594 FKM-ISMNHVYVNNLISRFKSKKCLXKSEWKNCFEHLVYEALKLAKDLSRTELMP 652  
 Qy 539 DADNSRERITLRFETLIDLVAHMPGCAVSKAEMLEKCRMIDLNG-----DQVLDN 589  
 RESULT 23  
 QAYOL8\_PLACH

Db 653 DKNDLGKVSSPFBQVLKDNLNINLSSAQVRLM---RMINNSLCNKNRVQENDKIDVA 708  
 Qy 590 BFLBAPR-----LSDLRK--EQQDENIRRS 614  
 Db 709 EFKGKMRVYCLAINKYINNEKVQLIETIGKHLADSSEIANHYRYREQEDENAVR 768  
 Qy 615 TGRPSVAKT 623  
 Db 769 RKKSSVIKS 777

RESULT 24  
 Q81728 TRYCR PRELIMINARY; PRT; 923 AA.  
 ID Q81728 TRYCR PRELIMINARY; PRT; 923 AA.  
 AC Q81728\_01-MAR-2003 (Tremblrel. 23, Created)  
 DT 01-MAR-2003 (Tremblrel. 23, Last sequence update)  
 RA 01-MAR-2004 (Tremblrel. 26, Last annotation update)  
 DE ICC31\_31.  
 Name=TCG31\_31;  
 OS Trypanosoma cruzi.  
 OC Schizotrypanum.  
 RN NCBI\_TaxID=5693;  
 RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN-Brener;  
 RX MEDLINE#98391765; PubMed=9724326;  
 RA Patterson B., Astlund L., Tammi M., Tran A.N., Hoheisel J.D.,  
 Peterson U., "Complete Sequence of a 93.4-kb contig from chromosome 3 of Trypanosoma cruzi containing a strand-switch region.", Genome Res. 8:809-816(1998).  
 [2]

NUCLEOTIDE SEQUENCE.

RA STRAIN-Brener,  
 Anderson B., Bontempi B.J., Submitted (JUN-2003) to the EMBL/GenBank/DDBJ databases.  
 DR EMBL: AC137088; AAN78342.1; -; Genomic\_DNA.  
 HSSP: P36873; 1176.  
 GO: GO:0016787; F:hydrolase activity; IEA.  
 DR InterPro, IPR01992; BP:Hand type; InterPro, IPR002048; EF-hand\_Ca\_bd.  
 DR InterPro, IPR004843; M-pesterase.  
 DR InterPro, IPR06186; T\_phnase\_apAH.  
 DR PF00449; Metallophos; 1.  
 DR PRINTS; PRO00114; STPHPTASE.  
 DR PRODom; P000252; Tphnase\_apAH; 1.  
 DR SMART; SW01536; PP2AC; 1.  
 DR PROSITE; PS00125; SER THR PHOSPHATASE; UNKNOWN 1.  
 SQ SEQUENCE 923 AA; 104798 MW; 956646BE32B2143 CRC64;

Query Match Best Local Similarity 17.3%; Score 594.5; DB 2; Length 923; Matches 174; Conservative 73; Mismatches 206; Indels 75; Gaps 15;

Qy 128 ILRRAAKSLQQLNNI---SPVSTAVS-----QQTVCGDGLKGKDDLLVILVKNGLP 176  
 DR 299 VFTDMASHMTMNVWRLSPPVGARVSVGRNGQSKVUVVGDLQALDLHILTKGCMP 358

Qy 177 SSNSPYVNGDFYDRGKGGLGEVLLILSLYLAPNAVLQRNGHEDSYNARYGTFREV 236  
 DR 359 NEGYTYIFNGDFVDRGANEGEVVILILPSIMLACKPKIVTNRGNHECDYMDNSYGFDEVs 418

Qy 237 SKYPRNHRKILARLDEVTRWMLPGISVLMNSRVLVHGGF--SOSTSLDIKSKIDRGKYSI 294  
 DR 419 TKYDRN---VFLVQRCFCALPLATIGKXPFVWGHGLPRKGXNIEDRISRQFRQI-- 473

Qy 295 LRPPLTPBRPLDKWENQOIF-DWMSDPOATMGCVNPNTLRGAGWMPGDYTMPLQHRL 353  
 DR 474 ---PMPD---YSQEBEDBIFODLWSDPVEDLOGWRSRPRGAGGVNGADVTBPLONGL 527

RESULT 25  
 Q812P5 ENTHI PRELIMINARY; PRT; 473 AA.  
 ID Q812P5 ENTHI PRELIMINARY; PRT; 473 AA.  
 AC Q812P5\_13-SEP-2005 (Tremblrel. 31, Created)  
 DT 13-SEP-2005 (Tremblrel. 31, Last sequence update)  
 RA 13-SEP-2005 (Tremblrel. 31, Last annotation update)  
 DE Protein phosphatase, putative.  
 GN ORFnames=90 t0028;  
 OS Entamoeba histolytica HM-1:IMSS.  
 OC Eukaryota; Entamoebidae; Entamoeba.  
 RN NCBI\_TaxID=29431;  
 RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN-HM-1:IMSS;  
 RX PubMed:15729342; DOI:10.1038/nature03291;  
 RA Loftus B., Anderson T., Davies R., Alsmark U.C., Samuelson J., Amodeo P., Roncaglia P., Berrianan M., Hirt R.P., Mann B.J., Nozaki T., Sun B., Pop M., Duchene M., Ackers J., Tannich B., Leippe M., Hofer M., Bruchhaus T., Wilhoft U., Bhattacharay A., Chillingworth T., Churcher C., Hance Z., Harris B., Harris D., Jagels K., Moule S., Mungall K., Ormond D., Squares R., Whitehead S., Quail M.A., Rabbinowitsch B., Norbertczak H., Price C., Wang Z., Guillen N., Gilchrist C., Stroup S.E., Bhattacharya S., Lohia A., Foster P.G., Sickeritz-Ponten T., Weber C., Singh U., Mukherjee C., El-Sayed N.M., Petrich W.A., Clark C.G., Embley T.M., Barrell B., Fraser C.M., Hall N.; "The genome of the protist parasite Entamoeba histolytica.", Nature 433:865-868 (2005).  
 -1 CAUTION: The sequence shown here is derived from an EMBL/Genbank/DDBJ whole genome shotgun (WGS) entry which is preliminary data.

CC DR EMBL: AAPB0100029; EAL4740\_1; -; Genomic DNA.  
 CC SEQUENCE 473 AA; 53606 MW; 665B460BCB0D9318 CRC64;

Query Match Best Local Similarity 16.9%; Score 582.5; DB 2; Length 473; Matches 143; Conservative 74; Mismatches 134; Indels 53; Gaps 14;

Qy 11 FLOWRKHQ-----REMRQRCWQIQONLBYASBDQABLYKEFDLKLKMPDAG 64  
 DR 68 FVKGKYMQASAFALGQLBATSA----BYAKKSPKD----GMNSMKGLK-- 115

Qy 65 RKNQYQSAHSVYLDKDDL---VERFGDIVAKTRPLTRKHDILIDVTRKKRN-- 118  
 DR 116 RKEQFPEA-SVYDEEQTISRDIDASTSTKIEDKPKITKENVQ--EYEAESNC 171

Qy 119 ---RFLKQYVLLIRAKSRSKOLPISPVSTAVSPOVTCGDLGKGLDQDLYVLRKG 174  
 DR 172 YGKCAIMLHKYCLRIERLSEAIKRNLYBETIQ-GRMETVGDHQFPDLHLFPRING 230

Qy 175 LOSSNPYVNGDFYDRGKGGLGEVLLILSLYLAPNAVLQRNGHEDSYNARYGTFREV 234

Db	231	LPSEDTYLFNGDPDVGSPGFCVCUTLFSPLIIVPNSVFLARGNHTRAMAMGPRGE	290
Qy	235	VESKPRMKRKRITLAFTDIFERYMLPLSVLNSRLVLTIVGGPSDS-TSDLIKSIDRKXVS	293
Db	291	VTKY--NENVTLAFSD-VFMOLPVCAVIDSKVVFVNGGIPPMYISLDIKKKR-----	343
Qy	294	ILRPLPLDEPKDKTRQFIDIMSPQATMGCVPLRGAGWFGPDVDMPLQRHRL	353
Db	344	-----CDPQEGSIASLWADPOTNSSSP-SIRGCGKSFQDITINPLDQNTL	391
Qy	354	SYVIRSHCKPNHEFMHDNKITIFSAASNLYIGNSKGAYIRL	397
Db	392	QYIVRSHEMKONGYEWGAGRLITVFSAPNYPOMNNKGAYHV	435
RESULT	26	QNE88_CAEEL PRELIMINARY;	PRT; 496 AA.
ID	QNE88	AC	QNE88;
AC	01-OCT-2000 (TREMBLrel. 15, Created)	DT	01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT	01-MAR-2004 (TREMBLrel. 26, Last annotation update)	DT	01-MAR-2004 (TREMBLrel. 26, Last annotation update)
DE	Hypothetical protein Y39B6A.2.	GN	ORFNames=Y39B6A.2;
OS	Caenorhabditis elegans.	OS	Caenorhabditis elegans.
OC	Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;	OC	Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC	Rhabditidae; Pelodoridae; Caenorhabditis.	OC	Rhabditidae; Pelodoridae; Caenorhabditis.
NCBI_TAXID=6239;	RN	[1]	
NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].	RP		
STRAIN=Bristol N2;	RX		
MEDLINE=98069613; PubMed=9851916;	RX		
The C. elegans sequencing consortium;	RT		
"Genome sequence of the nematode <i>C. elegans</i> : a platform for investigating biology.",	RT		
Science 282:2012-2018(1998).	RT		
PIR; T45088; T50401; TA17.	RL		
HSSP; P53041; TA17.	DR		
EMBL; AL122948; CAC51076.2; -; Genomic_DNA.	DR		
PIR; T45088.	DR		
PIR; T45088; T50401; TA17.	DR		
IntAct; QNE88; -;	DR		
Y39B6A.2; Caenorhabditis elegans.	DR		
Wormbase; WBGene00012665; Y39B6A.2.	DR		
WormPep; Y39B6A.2; CT36619.	DR		
Cytoplasm; IEA.	DR		
GO; GO:0005377; C:cytoplasm; IEA.	DR		
C:nucleus; IEA.	DR		
GO; GO:005634; C:nucleus; IEA.	DR		
F:binding; IEA.	DR		
GO; GO:0016787; F:hydrolyase activity; IEA.	DR		
GO; GO:004721; F:phosphoprotein phosphatase activity; IEA.	DR		
P:protein amino acid dephosphorylation; IEA.	DR		
GO; GO:0005470; P:protein amino acid dephosphorylation; IEA.	DR		
InterPro; IPR01483; M-peptidase.	DR		
InterPro; IPR01235; PPtase_5.	DR		
InterPro; IPR005186; T_phospe_apah.	DR		
InterPro; IPR01440; T_phospe_apah.	DR		
InterPro; IPR01990; TPR-like_helical.	DR		
Metallophos; 1.	DR		
PFam; PF00149; Metallophos; 1.	DR		
PRSR; PRSR03096; PPtase_5; 1.	DR		
PRINTS; PRO0114; STPHePTASE.	DR		
PRODOM; PD000252; T_phospe_apah; 1.	DR		
SMART; SM00022; TPR; 2.	DR		
PROSITE; PS50005; TPR; 2.	DR		
Complete proteome; Hypothetical Protein.	DR		
SEQUENCE 496 AA; 56462 MW; 2C872RF1700384EE CRC64;	SQ		
Query Match	16.9%	Score 581; DB 2; Length 496;	
Best Local Similarity	33.0%	Score 581; DB 2; Length 496;	
Matches	137;	Conservative 77; Mismatches 167; Indels 34; Gaps 12;	
Qy	3	ENAIRAAI--FTQKWRHQARREMR-RCNWQIQNLVEASQDQELYKF--FENDL	56
86	DNAI-AIDPSVYKGRKATAMALGSPFKKALTDYQAVWKWCNPNDKARAKFDECSKV	143	
RESULT	26	QNE88_CAEEL PRELIMINARY;	PRT; 829 AA.
ID	QNE88	AC	QNE88;
AC	01-OCT-2000 (TREMBLrel. 15, Created)	DT	01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT	01-MAR-2004 (TREMBLrel. 26, Last annotation update)	DT	01-MAR-2004 (TREMBLrel. 26, Last annotation update)
DE	Hypothetical protein Y39B6A.2.	GN	ORFNames=Y39B6A.2;
OS	Caenorhabditis elegans.	OS	Caenorhabditis elegans.
OC	Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;	OC	Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC	Rhabditidae; Pelodoridae; Caenorhabditis.	OC	Rhabditidae; Pelodoridae; Caenorhabditis.
NCBI_TAXID=6239;	RN	[1]	
NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].	RP		
STRAIN=Bristol N2;	RX		
MEDLINE=98069613; PubMed=9851916;	RX		
The C. elegans sequencing consortium;	RT		
"Genome sequence of the nematode <i>C. elegans</i> : a platform for investigating biology.",	RT		
Science 282:2012-2018(1998).	RT		
PIR; T45088; T50401; TA17.	RL		
HSSP; P53041; TA17.	DR		
EMBL; AL122948; CAC51076.2; -; Genomic_DNA.	DR		
PIR; T45088.	DR		
PIR; T45088; T50401; TA17.	DR		
IntAct; QNE88; -;	DR		
Y39B6A.2; Caenorhabditis elegans.	DR		
Wormbase; WBGene00012665; Y39B6A.2.	DR		
WormPep; Y39B6A.2; CT36619.	DR		
Cytoplasm; IEA.	DR		
GO; GO:0005377; C:cytoplasm; IEA.	DR		
C:nucleus; IEA.	DR		
GO; GO:005634; C:nucleus; IEA.	DR		
F:binding; IEA.	DR		
GO; GO:0016787; F:hydrolyase activity; IEA.	DR		
GO; GO:004721; F:phosphoprotein phosphatase activity; IEA.	DR		
P:protein amino acid dephosphorylation; IEA.	DR		
GO; GO:0005470; P:protein amino acid dephosphorylation; IEA.	DR		
InterPro; IPR01483; M-peptidase.	DR		
InterPro; IPR01235; PPtase_5.	DR		
InterPro; IPR005186; T_phospe_apah.	DR		
InterPro; IPR01440; T_phospe_apah.	DR		
InterPro; IPR01990; TPR-like_helical.	DR		
Metallophos; 1.	DR		
PFam; PF00149; Metallophos; 1.	DR		
PRSR; PRSR03096; PPtase_5; 1.	DR		
PRINTS; PRO0114; STPHePTASE.	DR		
PRODOM; PD000252; T_phospe_apah; 1.	DR		
SMART; SM00022; TPR; 2.	DR		
PROSITE; PS50005; TPR; 2.	DR		
Complete proteome; Hypothetical Protein.	DR		
SEQUENCE 496 AA; 56462 MW; 2C872RF1700384EE CRC64;	SQ		
Query Match	16.8%	Score 578; DB 2; Length 829;	
Best Local Similarity	33.0%	Score 578; DB 2; Length 829;	
Matches	137;	Conservative 77; Mismatches 167; Indels 34; Gaps 12;	
Qy	3	ENAIRAAI--FTQKWRHQARREMR-RCNWQIQNLVEASQDQELYKF--FENDL	56
86	DNAI-AIDPSVYKGRKATAMALGSPFKKALTDYQAVWKWCNPNDKARAKFDECSKV	143	
RESULT	27	QNE88_CAEEL PRELIMINARY;	PRT; 829 AA.
ID	QNE88	AC	QNE88;
AC	01-OCT-2000 (TREMBLrel. 15, Created)	DT	01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT	01-MAR-2004 (TREMBLrel. 31, Last annotation update)	DT	01-MAR-2004 (TREMBLrel. 31, Last annotation update)
DE	PPI-like protein serine/threonine phosphatase, putative	DB	(fragment);
OS	Plasmodium berghei.	OS	Plasmodium berghei.
OC	Bukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.	OC	Bukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
NCBI_TAXID=5821;	RN	[1]	
NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].	RP		
STRAIN=Bristol N2;	RX		
MEDLINE=98069613; PubMed=9851916;	RX		
The C. elegans sequencing consortium;	RT		
"Genome sequence of the nematode <i>C. elegans</i> : a platform for investigating biology.",	RT		
Science 282:2012-2018(1998).	RT		
PIR; T45088; T50401; TA17.	RL		
HSSP; P53041; TA17.	DR		
EMBL; AL122948; CAC51076.2; -; Genomic_DNA.	DR		
PIR; T45088.	DR		
PIR; T45088; T50401; TA17.	DR		
IntAct; QNE88; -;	DR		
Y39B6A.2; Caenorhabditis elegans.	DR		
Wormbase; WBGene00012665; Y39B6A.2.	DR		
WormPep; Y39B6A.2; CT36619.	DR		
Cytoplasm; IEA.	DR		
GO; GO:0005377; C:cytoplasm; IEA.	DR		
C:nucleus; IEA.	DR		
GO; GO:005634; C:nucleus; IEA.	DR		
F:binding; IEA.	DR		
GO; GO:0016787; F:hydrolyase activity; IEA.	DR		
GO; GO:004721; F:phosphoprotein phosphatase activity; IEA.	DR		
P:protein amino acid dephosphorylation; IEA.	DR		
GO; GO:0005470; P:protein amino acid dephosphorylation; IEA.	DR		
InterPro; IPR01483; M-peptidase.	DR		
InterPro; IPR01235; PPtase_5.	DR		
InterPro; IPR005186; T_phospe_apah.	DR		
InterPro; IPR01440; T_phospe_apah.	DR		
InterPro; IPR01990; TPR-like_helical.	DR		
Metallophos; 1.	DR		
PFam; PF00149; Metallophos; 1.	DR		
PRSR; PRSR03096; PPtase_5; 1.	DR		
PRINTS; PRO0114; STPHePTASE.	DR		
PRODOM; PD000252; T_phospe_apah; 1.	DR		
SMART; SM00022; TPR; 2.	DR		
PROSITE; PS50005; TPR; 2.	DR		
Complete proteome; Hypothetical Protein.	DR		
SEQUENCE 496 AA; 56462 MW; 2C872RF1700384EE CRC64;	SQ		
Query Match	16.9%	Score 581; DB 2; Length 496;	
Best Local Similarity	33.0%	Score 581; DB 2; Length 496;	
Matches	137;	Conservative 77; Mismatches 167; Indels 34; Gaps 12;	
Qy	3	ENAIRAAI--FTQKWRHQARREMR-RCNWQIQNLVEASQDQELYKF--FENDL	56
86	DNAI-AIDPSVYKGRKATAMALGSPFKKALTDYQAVWKWCNPNDKARAKFDECSKV	143	
RESULT	27	QNE88_CAEEL PRELIMINARY;	PRT; 829 AA.
ID	QNE88	AC	QNE88;
AC	01-OCT-2000 (TREMBLrel. 15, Created)	DT	01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT	01-MAR-2004 (TREMBLrel. 31, Last annotation update)	DT	01-MAR-2004 (TREMBLrel. 31, Last annotation update)
DE	PPI-like protein serine/threonine phosphatase, putative	DB	(fragment);
OS	Plasmodium berghei.	OS	Plasmodium berghei.
OC	Bukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.	OC	Bukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
NCBI_TAXID=5821;	RN	[1]	
NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].	RP		
STRAIN=Bristol N2;	RX		
MEDLINE=98069613; PubMed=9851916;	RX		
The C. elegans sequencing consortium;	RT		
"Genome sequence of the nematode <i>C. elegans</i> : a platform for investigating biology.",	RT		
Science 282:2012-2018(1998).	RT		
PIR; T45088; T50401; TA17.	RL		
HSSP; P53041; TA17.	DR		
EMBL; AL122948; CAC51076.2; -; Genomic_DNA.	DR		
PIR; T45088.	DR		
PIR; T45088; T50401; TA17.	DR		
IntAct; QNE88; -;	DR		
Y39B6A.2; Caenorhabditis elegans.	DR		
Wormbase; WBGene00012665; Y39B6A.2.	DR		
WormPep; Y39B6A.2; CT36619.	DR		
Cytoplasm; IEA.	DR		
GO; GO:0005377; C:cytoplasm; IEA.	DR		
C:nucleus; IEA.	DR		
GO; GO:005634; C:nucleus; IEA.	DR		
F:binding; IEA.	DR		
GO; GO:0016787; F:hydrolyase activity; IEA.	DR		
GO; GO:004721; F:phosphoprotein phosphatase activity; IEA.	DR		
P:protein amino acid dephosphorylation; IEA.	DR		
GO; GO:0005470; P:protein amino acid dephosphorylation; IEA.	DR		
InterPro; IPR01483; M-peptidase.	DR		
InterPro; IPR01235; PPtase_5.	DR		
InterPro; IPR005186; T_phospe_apah.	DR		
InterPro; IPR01440; T_phospe_apah.	DR		
InterPro; IPR01990; TPR-like_helical.	DR		
Metallophos; 1.	DR		
PFam; PF00149; Metallophos; 1.	DR		
PRSR; PRSR03096; PPtase_5; 1.	DR		
PRINTS; PRO0114; STPHePTASE.	DR		
PRODOM; PD000252; T_phospe_apah; 1.	DR		
SMART; SM00022; TPR; 2.	DR		
PROSITE; PS50005; TPR; 2.	DR		
Complete proteome; Hypothetical Protein.	DR		
SEQUENCE 496 AA; 56462 MW; 2C872RF1700384EE CRC64;	SQ		
Query Match	16.8%	Score 578; DB 2; Length 829;	
Best Local Similarity	33.0%	Score 578; DB 2; Length 829;	
Matches	137;	Conservative 77; Mismatches 167; Indels 34; Gaps 12;	
Qy	3	ENAIRAAI--FTQKWRHQARREMR-RCNWQIQNLVEASQDQELYKF--FENDL	56
86	DNAI-AIDPSVYKGRKATAMALGSPFKKALTDYQAVWKWCNPNDKARAKFDECSKV	143	

**Matches** 179; **Conservative** 108; **Mismatches** 247; **Indels** 120; **Gaps** 18;

**Qy** 63 AGRKNNQYQGSARHVSVLDSKDLYVERGIVNAK--IR---IPRKWHDLLIDVFRKKR 116  
**DR** DR InterPro; IPR004843; M-peptidase.

**Db** 27 SSNSNQISGSGTISYQMSYDSNPKLKKVKNRKPAEFLNPLLTQNTVPLSMWYK-- 83  
**DR** DR InterPro; IPR001236; PPtase\_5.

**Qy** 117 GNRILHPKVVALIIRBAAKSILQ---LPNISPVSTAVSQVTVGCGDQLRGKLUDDLVVLH 171  
**DR** DR InterPro; IPR001440; TPR.

**Db** 84 -----ILNETKOMLSERIKSSVLLNDITTKSKDTKLVLGDVHGQLANDVNLFN 132  
**DR** DR InterPro; IPR00186; T phase\_apAH.

**Qy** 172 KNGLIPSSSNPYPVENGDFTRGKGLEVLILLSLYLAPNAAVFLNQRHEDSYMMARYGP 231  
**DR** DR Pfam; PF00149; Metallophos; 1.

**Db** 133 RFLGLPSSNNIYIYINGDIADRGGONABIFPLFAFKLSNVDSVLRNGHICCSYNEYGP 192  
**DR** DR Pfam; PF00515; TPR\_1; 3.

**Qy** 232 IREVESKYPNHRGILATIDBVRLPLIGSVLNSRLVHGGSDSTSLSLDL--IKSDRG 289  
**DR** DR Prints; PR00114; STPHTAAB.

**Db** 193 HNRKVLTSKYDSE--VFDFIOFEIFELLSLSVNQNOIPTVHGGSLRYPDLSVQDIDKDR 249  
**DR** DR SMART; SM01356; PP2AC; 1.

**Qy** 290 KYVSLIRPPLTDGEPLDKTYEWQOIFDIMSDPOATMGCVPTNLRGAGWFGPDTVDNPLQ 349  
**DR** DR SMART; SM00228; TPR; 3.

**Db** 250 KH-BILHP----EKVDETTIFDLWNSDQXDGKIGGGARGNCIAFGFDVTEFLK 300  
**DR** DR PROSITE; PS50005; TPR; 3.

**Qy** 350 RHLRLSWSVRSHCKP--NGHEFMDNKITITASASNYAIGSNGKAYTRLNQMLPHVQ 407  
**DR** DR PROSITE; PS50293; TPR\_REGION; 2.

**Db** 301 KNKEDIIITRSHQVPTKLGEIESHEKGKCTILFSASNYCNKTKNLAATIFNQDLTVEEVH 360  
**DR** DR SQ Sequence; 594 AA; 69196 MW; 63B838750664181C CRC64;

**Qy** 408 YISAASQYKRLS----KQRMGVESALKE----- 439  
**DR** DR Query Match 16.8%; Score 577.5; DB 2; Length 594;  
**Db** 361 YMSPLSDVIRETFEBENOKLBRKLYNSNILELEKNEORNSKQIPPEGALDTINPESTLI 420  
**DR** DR Best Local Similarity 35.9%; Pred. No. 1..3e-26;  
**Qy** 440 RDRHDELEDEFRKYDPKGPSQYVISHWCKVMMENVTKL-GLPNRPLRDRGKLAQPGTDSQKVN 498  
**DR** DR Matches 151; Conservative 77; Mismatches 144; Indels 49; Gaps 17;

**Db** 421 CMEKQNLWNLVYKDKENTGKHMINTWRBEELEKLTQVPMWYLCKKFM-IENNHYV 479  
**DR** DR 184 IDBAIKNPYY---AKAYRKGCSYNLLSLSLKRASCFCQVKLTDKNSBLK-LIQCK 238

**Qy** 499 NRRTDLDJTD----VILAEADG--MSVMDDLYANAKSAYLAIFIENDTADNSGIRTLDE 550  
**DR** DR 64 GR--KNQYQGSARHVSVDK---DDIVERGEDIWAKIELP--RNKHIDLLIDVFRKKR 116

**Db** 480 NNVLSRFKKINYDQGOTFLKESWNKNECFEHLIYEALKADLSRRETLMIFDKNLGKVSPSE 539  
**DR** DR 239 KLIPIFQPOKAIB--I-EKOMPYETTIVLDSLKLEN--MEAPTYDRNNTNL--DFLKCKVA 291

**Qy** 551 FEAIDLIVAHMGGAYSAEML---EKCRMADLINGDKVDLNPLRFAFR----- 596  
**DR** DR 352 KINGYPSBKNSYLFNGDFVDRGGSFSVWVIFLYLAKUTFPNNVVLITGNHETDNMKYIG 411

**Db** 540 FEGVQLDQNIINSAEBOVIRLMLLINSNLSCKNKRQIENDKDVIAERTGKMRVYCLAINK 599  
**DR** DR 291 RIREVESKYPNHRKVALIIRBAAKSILQDLYVERGIVNAK--IR---IPRKWHDLLIDVFRKKR 288

**Db** 597 -----LSDLHKR--EQDENIRRSTARSPSWAKT 623  
**DR** DR 171 HNRKLVKVALIIRBAAKSILQDLYVERGIVNAK--IR---IPRKWHDLLIDVFRKKR 63  
**Db** 600 BYINNAKVKQKLIETIGKHLADSSEBIANFHRYFEEQDESENDRRKRSIVKS 653  
**DR** DR 412 FLGBLOSKY--DEBKMHVFLPSDS-FKPLPLAVYLNKNTIPICHGQPSKSTDTLBDIKDR 468

**Qy** 408 Y 408  
**DR** DR 289 GKYSVSLRPLTDGEPLDKTYEWQOIFDIMSDPOATMGCVPTNLRGAGWFGPDTVDNPLQ 348

**Db** 469 NK-----EPDE--GWMDDLWSDPBNPBBEGPKSK-RGJGSGCTDIRENPL 512  
**DR** DR 469 NK-----EPDE--GWMDDLWSDPBNPBBEGPKSK-RGJGSGCTDIRENPL 512

**Qy** 349 QHLRLSWSVRSHCKP--NGHEFMDNKITITASASNYAIGSNGKAYTRL-NNQMLPHVQ 407  
**DR** DR 513 KINNLSLIRSRSHVEDREGSLEQNGQIYTFSAPNQCDIMKNGKAFLKPKGSNSIKPCV 572

**Db** 573 P 573

**RESULT** 28

**O8WQ3** PLAFA  
**ID** O8WQ3\_PLAFA PRELIMINARY; **PRT**; 594 AA.

**AC** O8WQ3,  
**DT** 01-MAR-2002 (Tremblrel. 20, Created)  
**DT** 01-MAR-2002 (Tremblrel. 20, Last sequence update)  
**DT** 01-MAR-2004 (Tremblrel. 26, Last annotation update)  
**DE** Serine/threonine protein phosphatase PP5.

**OS** Plasmodium falciparum.

**OC** Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.

**NCBI\_TaxID**=5833;

**[1]** NUCLEOTIDE SEQUENCE.  
**Medline**=2229045; PubMed=173764; DOI=10.1186/1471-2180-1-31;  
**RA** Dobson S, Kar B, Kumar R, Adams B, Barik S;  
**RT** "A novel tetra-tricopeptide repeat (TPR) containing PP5 serine/threonine protein phosphatase in the malaria parasite,"  
**RT** Plasmodium falciparum.;"  
**RT** BMC Microbiol, 1,3:1-3 (2001).  
**DR** EMBL; AY054983; ARNL15170.1; -; mRNA.  
**DR** HSSP; P36873; ITI6.  
**DR** GO; GO:0005737; C:cytoplasm; IBA.  
**DR** GO; GO:005334; C:nucleus; IBA.  
**DR** GO; GO:0016787; F:hydrolase activity; IBA.  
**DR** GO; GO:0004721; F:phosphoprotein phosphatase activity; IBA.  
**DR** GO; GO:0006470; P:protein amino acid dephosphorylation; IBA.

**RESULT** 29

**O8IDE7** PLAFT

**ID** O8IDE7\_PLAFT PRELIMINARY; **PRT**; 658 AA.

**AC** O8IDE7,  
**DT** 01-MAR-2003 (Tremblrel. 23, Created)  
**DT** 01-MAR-2003 (Tremblrel. 23, Last sequence update)  
**DT** 01-MAR-2004 (Tremblrel. 23, Last annotation update)

**DB** Serine/threonine protein phosphatase pfp5 (SC 3.1.3.16).  
**GN** Name=pfp5; synonym=Malu13P1.274;  
**OS** Plasmodium falciparum (isolate 3D7);  
**OC** Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.

**NCBI\_TaxID**=36329;

**[1]** NUCLEOTIDE SEQUENCE.  
**RC** STRAIN-3D7,  
**RA** Harris B, Lennard N, Clark L, Barron A, Corton C,  
**RA** Berriman M, Pain A, Hall N, Atkin R, Chillingworth C, Doggett J,,  
**RA** Ormond D., Salter M., Haynes R., Hall S., Quail M., Barrell B.,  
**RA** Submitted (SBP-2002) to the EMBL/GenBank/DBJ databases.  
**DR** EMBL; ALB44509; CAD2675.1; -; Genomic\_DNA.  
**DR** HSSP; P36873; ITI6.  
**GO**; GO:0005737; C:cytoplasm; IBA.



RA	Hradecity P., Huang Y., Kaminker J.S., Millburn G.H., Prochnik S.E., Smith C.D., Tupy J.L., Whirfield S.J., Bayraktaroglu L., Berman B.P., Bettencourt R.R., Celmer S.E., de Grey A.N.J., Drisdale R.A., Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,' Lewis S.B.; Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,	Smith C.D., Tupy J.L., Whirfield S.J., Bayraktaroglu L., Berman B.P., Bettencourt R.R., Celmer S.E., de Grey A.N.J., Drisdale R.A., Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,' Lewis S.B.; Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,	Smith C.D., Tupy J.L., Whirfield S.J., Bayraktaroglu L., Berman B.P., Bettencourt R.R., Celmer S.E., de Grey A.N.J., Drisdale R.A., Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,' Lewis S.B.; Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
RT	"Annotation of the <i>Drosophila melanogaster</i> euchromatic genome: a systematic review"; Genome Biol. 3:RESEARCH0083.1-RESEARCH0083.22(2002).	"Annotation of the <i>Drosophila melanogaster</i> euchromatic genome: a systematic review"; Genome Biol. 3:RESEARCH0083.1-RESEARCH0083.22(2002).	"Annotation of the <i>Drosophila melanogaster</i> euchromatic genome: a systematic review"; Genome Biol. 3:RESEARCH0083.1-RESEARCH0083.22(2002).
RN	[6]	[6]	[6]
RP	NUCLEOTIDE SEQUENCE.	NUCLEOTIDE SEQUENCE.	NUCLEOTIDE SEQUENCE.
RG	Berkley Drosophila Genome Project;	"Drosophila melanogaster release 4 sequence.";	"Drosophila melanogaster release 4 sequence.";
RA	Celniker S., Carlson J., Wan K., Peiffer B., Frise E., George R., Hobkins R., Rubin G., Stapleton M., Paclob J., Park S., Svirkas R., Smith E., Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.	Celniker S., Carlson J., Wan K., Peiffer B., Frise E., George R., Hobkins R., Rubin G., Stapleton M., Paclob J., Park S., Svirkas R., Smith E., Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.	Celniker S., Carlson J., Wan K., Peiffer B., Frise E., George R., Hobkins R., Rubin G., Stapleton M., Paclob J., Park S., Svirkas R., Smith E., Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
RL	NUCLEOTIDE SEQUENCE.	NUCLEOTIDE SEQUENCE.	NUCLEOTIDE SEQUENCE.
RN	Submitted (MAR-2005) to the EMBL/GenBank/DBJ databases.	Submitted (MAR-2005) to the EMBL/GenBank/DBJ databases.	Submitted (MAR-2005) to the EMBL/GenBank/DBJ databases.
RP	[8]	[8]	[8]
RC	STRAIN=Berkeley;	STRAIN=Berkeley;	STRAIN=Berkeley;
RA	Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J., Champine M., Chavez C., Dorsett V., Faafan D., Frise E., George R., Gonzalez M., Guarin H., Li P., Liao G., Miranda A., Mungall C.J., Nunojo J., Paclob J., Paragab V., Park S., Phouanenavong S., Wan K., Yu C., Lewis S.B., Rubin G.M., Celniker S., Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.	Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J., Champine M., Chavez C., Dorsett V., Faafan D., Frise E., George R., Gonzalez M., Guarin H., Li P., Liao G., Miranda A., Mungall C.J., Nunojo J., Paclob J., Paragab V., Park S., Phouanenavong S., Wan K., Yu C., Lewis S.B., Rubin G.M., Celniker S., Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.	Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J., Champine M., Chavez C., Dorsett V., Faafan D., Frise E., George R., Gonzalez M., Guarin H., Li P., Liao G., Miranda A., Mungall C.J., Nunojo J., Paclob J., Paragab V., Park S., Phouanenavong S., Wan K., Yu C., Lewis S.B., Rubin G.M., Celniker S., Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
RL	EMBL; AE003684; AAP54438.1; -; Genomic_DNA.	EMBL; AE003684; AAP54438.1; -; mRNA.	EMBL; AE003684; AAP54438.1; -; mRNA.
RN	EMBL; AA058356; AAL13585.1; -; mRNA.	EMBL; AA058356; AAL13585.1; -; mRNA.	EMBL; AA058356; AAL13585.1; -; mRNA.
RP	NUCLEOTIDE SEQUENCE.	NUCLEOTIDE SEQUENCE.	NUCLEOTIDE SEQUENCE.
RC	RESULT 31	RESULT 31	RESULT 31
RA	OSCJAB_CRYO	OSCJAB_CRYO	OSCJAB_CRYO
DR	OSCJAB_CRYO PRELIMINARY;	OSCJAB_CRYO PRELIMINARY;	OSCJAB_CRYO PRELIMINARY;
AC	PRT; 525 AA.	PRT; 525 AA.	PRT; 525 AA.
DT	05-MAY-2005 (TREMBLrel. 30, Created)	05-MAY-2005 (TREMBLrel. 30, Last sequence update)	05-MAY-2005 (TREMBLrel. 30, Last annotation update)
DT	10-MAY-2005 (TREMBLrel. 30, Last sequence update)	10-MAY-2005 (TREMBLrel. 30, Last annotation update)	10-MAY-2005 (TREMBLrel. 30, Last annotation update)
DR	Phosphoprotein phosphatase-related.	Phosphoprotein phosphatase-related.	Phosphoprotein phosphatase-related.
GN	ORFNAME=Cryo.20310;	ORFNAME=Cryo.20310;	ORFNAME=Cryo.20310;
OS	Cryptosporidium hominis.	Cryptosporidium hominis.	Cryptosporidium hominis.
OC	Cryptosporidiidae; Apicomplexa; Coccidia; Bimerida;	Cryptosporidiidae; Apicomplexa; Coccidia; Bimerida;	Cryptosporidiidae; Apicomplexa; Coccidia; Bimerida;
OX	NCBI_TAXID=237895;	NCBI_TAXID=237895;	NCBI_TAXID=237895;
RN	[1]	[1]	[1]
RP	NUCLEOTIDE SEQUENCE.	NUCLEOTIDE SEQUENCE.	NUCLEOTIDE SEQUENCE.
RC	STRAIN=TU502;	STRAIN=TU502;	STRAIN=TU502;
RA	Xu P., Widmer G., Wang Y., Ozaki L.S., Alves J.M., Serrano M.G., Ruiz D., Manque P., Akiyoshi D., Mackey A.J., Pearson W.R., Dear P.H., Bankier A.T., Peterson D.L., Abrahamson M.S., Kapur V., Tziori S., Buck G.A.; "The genome of <i>Cryptosporidium hominis</i> "; Nature 431:1107-1112(2004); EMBL; AB01000136; EAL13686.1; -; Genomic_DNA.	Xu P., Widmer G., Wang Y., Ozaki L.S., Alves J.M., Serrano M.G., Ruiz D., Manque P., Akiyoshi D., Mackey A.J., Pearson W.R., Dear P.H., Bankier A.T., Peterson D.L., Abrahamson M.S., Kapur V., Tziori S., Buck G.A.; "The genome of <i>Cryptosporidium hominis</i> "; Nature 431:1107-1112(2004); EMBL; AB01000136; EAL13686.1; -; Genomic_DNA.	Xu P., Widmer G., Wang Y., Ozaki L.S., Alves J.M., Serrano M.G., Ruiz D., Manque P., Akiyoshi D., Mackey A.J., Pearson W.R., Dear P.H., Bankier A.T., Peterson D.L., Abrahamson M.S., Kapur V., Tziori S., Buck G.A.; "The genome of <i>Cryptosporidium hominis</i> "; Nature 431:1107-1112(2004); EMBL; AB01000136; EAL13686.1; -; Genomic_DNA.
DR	InterPro; IPR00843; M-peptidase.	InterPro; IPR00843; M-peptidase.	InterPro; IPR00843; M-peptidase.
DR	GO; GO_0016787; F:hydrolase activity; IEA.	GO; GO_0016787; F:hydrolase activity; IEA.	GO; GO_0016787; F:hydrolase activity; IEA.
DR	GO; GO_004721; P:phosphoprotein phosphatase activity; IEA.	GO; GO_004721; P:phosphoprotein phosphatase activity; IEA.	GO; GO_004721; P:phosphoprotein phosphatase activity; IEA.
DR	GO; GO_005470; P:protein amino acid dephosphorylation; IEA.	GO; GO_005470; P:protein amino acid dephosphorylation; IEA.	GO; GO_005470; P:protein amino acid dephosphorylation; IEA.
DR	InterPro; IPR011990; TPR-like helical.	InterPro; IPR011990; TPR-like helical.	InterPro; IPR011990; TPR-like helical.
DR	InterPro; IPR016186; T_phosphopeptide.	InterPro; IPR016186; T_phosphopeptide.	InterPro; IPR016186; T_phosphopeptide.
DR	PFAM; PF00149; Metallophos; 1.	PFAM; PF00149; Metallophos; 1.	PFAM; PF00149; Metallophos; 1.
DR	PFAM; PF00149; PPtase_5; 1.	PFAM; PF00149; PPtase_5; 1.	PFAM; PF00149; PPtase_5; 1.
DR	PRINTS; PRO0114; STPHPTASE.	PRINTS; PRO0114; STPHPTASE.	PRINTS; PRO0114; STPHPTASE.
DR	SMART; SM00156; PP2Ac; 1.	SMART; SM00156; PP2Ac; 1.	SMART; SM00156; PP2Ac; 1.
DR	InterPro; IPR006022; TPR; 3.	InterPro; IPR006022; TPR; 3.	InterPro; IPR006022; TPR; 3.
DR	PROSITE; PS500293; TPR_REGION; 1.	PROSITE; PS500293; TPR_REGION; 1.	PROSITE; PS500293; TPR_REGION; 1.
DR	KW_Hydrolase.	KW_Hydrolase.	KW_Hydrolase.
SQ	SEQUENCE 520 AA; 59263 MW; 47193DD9BDC38F32 CRC64;	SEQUENCE 520 AA; 59263 MW; 47193DD9BDC38F32 CRC64;	SEQUENCE 520 AA; 59263 MW; 47193DD9BDC38F32 CRC64;
Query Match	16.8%; Score 577; DB 2; Length 520;	16.7%; Score 575.5; DB 2; Length 525;	16.7%; Score 575.5; DB 2; Length 525;
Best Local Similarity	32.1%; Pred. No. 1.e-26;	34.7%; Pred. No. 1.e-26;	34.7%; Pred. No. 1.e-26;
Matches	136; Conservative 83; Mismatches 153; Indels 52; Gaps 15;	151; Conservative 67; Mismatches 160; Indels 57; Gaps 13;	151; Conservative 67; Mismatches 160; Indels 57; Gaps 13;
Qy	4 NAIRAAIFIQKWRHQAARREM--ORCNWQIQFQNLEYASBDQDLYKFENDIK- 57	4 NAIRAAIFIQKWRHQAARREM--ORCNWQIQFQNLEYASBDQDLYKFENDIK- 57	4 NAIRAAIFIQKWRHQAARREM--ORCNWQIQFQNLEYASBDQDLYKFENDIK- 57
Db	58 -----HMPRODGAKRKNYOCDSAHSVLDL-KDIDVPEFGDIVAKIEPDKRKHIDL 107	108 SAVKDPAVKGTYRRAAHMSLGKFQALCPERVKCR_PNDKAQKPFTECNKIVRM 166	108 LIDVFKRKGRLHPLPKVALLRRAASLQKQPNISPVSTAVSQVTVCGDLHGKDDLL 167
Qy	167 RAFFERIAIAVADKPERK-LSEMYSMENTEDDYKGPQED----GKVTLKPMKE-- 215	167 RAFFERIAIAVADKPERK-LSEMYSMENTEDDYKGPQED----GKVTLKPMKE-- 215	167 RAFFERIAIAVADKPERK-LSEMYSMENTEDDYKGPQED----GKVTLKPMKE-- 215
Db	89 FSKAYTRGIANFNLKYSLARKDFMVLNLQTQNDRAQSKIQICTKLIQKBFMDAIST 148	89 FSKAYTRGIANFNLKYSLARKDFMVLNLQTQNDRAQSKIQICTKLIQKBFMDAIST 148	89 FSKAYTRGIANFNLKYSLARKDFMVLNLQTQNDRAQSKIQICTKLIQKBFMDAIST 148

QY 50 ---KFENDLKKHNPQAAGRKNNQQGSAHVSVJDDKDLVERPGDIVAKI----ELP- 99  
 Db 149 DRSKLHLHTIDTMVAPGFASVNDYSG-PHYKPLVDHSRFAQS-GNETNLKSRPVVKPD 206  
 QY 100 ---IRKNHIDLILDEFRERKGRHLHPKVVALIREAKSKLKLPLNISPVSTAVSQQ--VT 154  
 Db 207 ASLJRESFVTELLD-FLKNPENRLHRKYAYMIVYDQVLQVKEVASKPLVRINIGKOBHT 265  
 QY 155 VCDPLHAKUDILVULHENGLOSSSNPYVNGDFVPGKGRGLFVLLSLYLAFNAPV 214  
 Db 266 VCDQDINGQPFDLNITFDINGLPSVNNGYLPGDFVDRGSFSVSEVILVLTPLTAKMYPYHV 325  
 QY 215 LNRGHEDSVMMARYGFTREVEKYPRHKRILAFIDEVYRMLPLGSVLSNRVLTIVGGP 274  
 Db 326 LARGNHETKLNLYKGEGEGRVAKYDSG--LYDLSBACTPLAHVINDKUPVHGL 382  
 QY 275 SPSTSLDLKKISDRGKYSVSLRPLPDTQEPDPLKTSMOIFDMSDPAOTMGCVPTLRG 334  
 Db 383 CSBDNVKL-----SDISOLYSRCPADS---GFMSSLWSDPOQKGRSPB-RG 428  
 QY 335 AGTMWFGPVTDFNLTQRLSLTVRSHECKNGHEPMNDKNTITFSAINTAIGSNKGAY 394  
 Db 429 VACNFGPDPVTLNFLKTNNLDYIIRSHEVTKOGSYVSDHDGCKCTVSPAPNYCDSMGNGAP 488  
 QY 395 IRLAN-QLMPHFWQY 408  
 Db 489 IKTHRYDVKPNFVQF 503

**RESULT 32**  
 Q4UFS6 \_THRAAN PRELIMINARY; PRT; 516 AA.  
 ID Q4UFS6 \_THRAAN PRELIMINARY; PRT; 516 AA.

AC 04UFS6\_ THRAAN PRELIMINARY, PRT; 516 AA.  
 DT 13-SEP-2005 (TREMBLrel. 31, Last sequence update)  
 DT 13-SEP-2005 (TREMBLrel. 31, Last annotation update)  
 Seine/threonine protein phosphatase, putative (EC 3.1.3.16).  
 ORFNames=TA5975;  
 OS Theileria annulata.  
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.  
 OC Theileria.  
 OC NCBI\_TaxID=5833;  
 RN [1] NUCLEOTIDE SEQUENCE.  
 RX MEDLINE-2189623; PubMed-11897131; DOI-10.1016/S0166-6851(02)00007-5;  
 RX Lindenthal C., Klinert M.Q.;  
 RT "Identification and biochemical characterisation of a protein phosphatase 5 homologue from Plasmodium falciparum.;"  
 RT Mol. Biochem. Parasitol. 120:257-268 (2002).  
 DR EMBL; AF404815; AAC05648.1; -; Genomic\_DNA.  
 DR HSSP; P36073; 1IT6.  
 DR GO; GO:0005737; C:cytoplasm; IEA.  
 DR GO; GO:0005634; C:nucleus; IEA.  
 DR GO; GO:0016787; F:hydrolase activity; IEA.  
 DR GO; GO:0004721; F:phosphoprotein phosphatase activity; IEA.  
 DR GO; GO:0006470; P:protein amino acid dephosphorylation; IEA.  
 DR InterPro; IPR004843; M-peptidase.  
 DR InterPro; IPR011236; PPase\_5.  
 DR InterPro; IPR001140; T\_Ptase\_apah.  
 DR Pfam; PF00149; Metallophosph; 1.  
 DR Pfam; PF00515; TPR\_1; 3.  
 DR PIRSF; PIRSF031095; PPase\_5; 1.  
 DR PRINTS; PRO0114; STPMPHASE.  
 DR SMART; SMD00156; PP2Ac; 1.  
 DR SMART; SMD0028; TPR; 3.  
 DR PROSITE; PS50005; TPR; 3.  
 DR PROSITE; PS55023; TPR REGION; 2.  
 DR PROSITE; PS55023; TPR REGION; 2.  
 SQ SEQUENCE 516 AA; 59193 MW; C670662EBFC4FA55 CRC64;

Query Match 16.7%; Score 574.5; DB 2; Length 516;  
 Best Local Similarity 38.2%; Pred. No. 1.6e-26;  
 Matches 120; Conservative 65; Mismatches 102; Indels 27; Gaps 8;

QY 103 NHID---LLIDV---FRKRGGRGKRLHPKTYVALLRBAAKSLKLQPNISPVSTAVSQQTVCQ 157  
 Db 200 NHHDKSYVFLHNTLBSYLVKPGNTTHKCYCMLLIVKTIRESVWLUNIYPCDELTVCQ 259  
 QY 158 DIAKGKDLDDLVVLAHKNGLPSSSNPYVNGDFVDRGKGRGLEVLLSLYLAPPNAVFLNR 217  
 Db 260 DINGQPFPLNITSINGPDTDENSYLFGDFVDRGSFSPECUTLFLAKVLPVPPSSFHVR 319  
 QY 218 GHNFEDSVNARVGRFIREVSKYPRHRKGLAEFLDEVRVRLPLGSVNSRVLIVHGRP-S 275  
 Db 320 GHNFEDSVNARVGRFIREVSKYPRHRKGLAEFLDEVRVRLPLGSVNSRVLIVHGRP-S 275  
 QY 276 DSLSLDLKKISDRGKYSVSLRPLPDTQEPDPLKTSMOIFDMSDPAOTMGCVPTLRG 335  
 Db 377 ENTYDJKKDFPK-----EPSD-----GJMTDLMWSDPEPSNLSPLS-T-RGV 420

**RESULT 33**  
 Q4UFS6\_ PLARA PRELIMINARY, PRT; 594 AA.  
 ID Q4UFS6\_ PLARA PRELIMINARY, PRT; 594 AA.

AC 04UFS6\_ PLARA PRELIMINARY, PRT; 594 AA.  
 DT 01-DEC-2001 (TREMBLrel. 19, Created)  
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)  
 DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)

DR Serine/threonine protein phosphatase PEPP5.  
 NamePEP5;  
 OS Plasmodium falciparum.  
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.  
 OC NCBI\_TaxID=5833;  
 RN [1] NUCLEOTIDE SEQUENCE.  
 RX MEDLINE-2189623; PubMed-11897131; DOI-10.1016/S0166-6851(02)00007-5;  
 RX Lindenthal C., Klinert M.Q.;  
 RT "Identification and biochemical characterisation of a protein phosphatase 5 homologue from Plasmodium falciparum.;"  
 RT Mol. Biochem. Parasitol. 120:257-268 (2002).  
 DR EMBL; AF404815; AAC05648.1; -; Genomic\_DNA.  
 DR HSSP; P36073; 1IT6.  
 DR GO; GO:0005737; C:cytoplasm; IEA.  
 DR GO; GO:0005634; C:nucleus; IEA.  
 DR GO; GO:0016787; F:hydrolase activity; IEA.  
 DR GO; GO:0004721; F:phosphoprotein phosphatase activity; IEA.  
 DR GO; GO:0006470; P:protein amino acid dephosphorylation; IEA.  
 DR InterPro; IPR004843; M-peptidase.  
 DR InterPro; IPR011236; PPase\_5.  
 DR InterPro; IPR001140; T\_Ptase\_apah.  
 DR Pfam; PF00149; Metallophosph; 1.  
 DR Pfam; PF00515; TPR\_1; 3.  
 DR PIRSF; PIRSF031095; PPase\_5; 1.  
 DR PRINTS; PRO0114; STPMPHASE.  
 DR SMART; SMD00156; PP2Ac; 1.  
 DR SMART; SMD0028; TPR; 3.  
 DR PROSITE; PS50005; TPR; 3.  
 DR PROSITE; PS55023; TPR REGION; 2.  
 SQ SEQUENCE 594 AA; 69248 MW; 64CA9D0004DBREFE CRC64;

Query Match 16.7%; Score 573.5; DB 2; Length 594;  
 Best Local Similarity 35.9%; Pred. No. 2.3e-26;  
 Matches 151; Conservative 77; Mismatches 144; Indels 49; Gaps 17;

QY 6 IRAAIITQKMRHBRMRMQRCCWQIQMLYEAS-E-QDQEALYKFFNDLKKRQA 63  
 Db 184 IDRAIKHNPKYY---AKAYTRKGCSYLLSDKLKRASCFKVKLTKDKNSELK-LKQCK 238  
 QY 64 GR---KNOYQGSAHVSVLDDK---DDLVESFGDIVAKIBIPI-RKHNHIDLIVFRKGR 116  
 Db 239 KLIPEQDQPKAIE---LHQRPYVETVLSKIN---MEAPIYRNMLI---DEPKVA 291  
 QY 117 -----GNRLHPKTYVALLRRAKSXKQLENISPVSTAVSQQTVCQDGKQDPLVLUV 170  
 Db 292 DVISIPNPKLNKCVCAIVLIVVLUKLPVLYMLBEDLTICGDVHQYYQDNLIM 351  
 QY 171 HKNGLSSSNPQYVNGDFVDRGKGRGLEVLLSLYLAPPNAVFLRGNHEDSVNARYG 230  
 Db 352 KINGYPSBKSYLFGDFVDRGSFSPECUTLFLAKVLPVPPSSFHVR 411  
 QY 231 FIREVSKYPRHRKGLAEFLDEVRVRLPLGSVNSRVLIVHGRP---SDSLSLTKSID 288  
 Db 412 FGEGLQKRY---DEKOMVLFSDS-PKELPLAVLNTNFIICHGIGSKTDTLSDERKID 468

Db	469	IK-----ELIDE---GUMDLWSDPNEKGPKESK-KGIGPSGTDITENPL	512	289 GKVSLRPLTGDPLDKTEWQDIDMWSPOATMGCVPTLURGAGWMFGPDYTDNF	348	Qy
Qy	349	ORHLSVIRSHBCKNGHEFMHDNKTITFSASNNTAIGSNKGAYTRL-NNQLMPHFVQ	407	349 ORHLSVIRSHBCKNGHEFMHDNKTITFSASNNTAIGSNKGAYTRL-NNQLMPHFVQ	407	Qy
Db	513	KINNLILIRSHBVRREGSISQNGQIYTFSAPNCDIMNKNGARLKPKNSIKEC	572	513 KINNLILIRSHBVRREGSISQNGQIYTFSAPNCDIMNKNGARLKPKNSIKEC	572	Db
Qy	408	Y 408	573	RESULT 34 Q5W6J3 ORYSA PRELIMINARY; PRT; 483 AA.	573 F 573	Qy
Db	573	Q5W6J3; 01-FEB-2005 (TREMBLrel. 29, Created) 01-FEB-2005 (TREMBLrel. 29, Last sequence update) 01-FEB-2005 (TREMBLrel. 29, Last annotation update)	573	01-FEB-2005 (TREMBLrel. 29, Last sequence update) 01-FEB-2005 (TREMBLrel. 29, Last annotation update)	573	Db
Qy	408	Putative serine/threonine phosphatase. Name=OSJNB0115P21; 05W6J3;	408	Putative serine/threonine phosphatase. Name=OSJNB0115P21; 05W6J3;	408	Qy
Db	573	Oryza sativa (Iaponica cultivar-group). Bukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzeae; Oryza. NCBI_TaxID=39947; [1]	573	Oryza sativa (Iaponica cultivar-group). Bukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzeae; Oryza. NCBI_TaxID=39947; [1]	573	Db
Qy	408	NUCLEOTIDE SEQUENCE. Chow T.-Y., Hsing Y.-I.C., Chen C.-S., Chen H.-H., Liu S.-M., Chen Y.-T., Chang S.-J., Chen H.-C., Chen S.-K., Chen T.-R., Chen Y.-L., Cheng C.-H., Chung C.-I., Han S.-Y., Hsiao S.-H., Heiung J.-N., Hsu C.-H., Huang J.-J., Kan P.-I., Lee M.-C., Leu H.-L., Li Y.-F., Lin S.-J., Lin Y.-C., Wu S.-W., Yu C.-Y., Yu S.-W., Wu H.-P., Shaw J.-F., Submitted (NOV-2004) to the EMBL/GenBank/DBJ databases.	408	NUCLEOTIDE SEQUENCE. Chow T.-Y., Hsing Y.-I.C., Chen C.-S., Chen H.-H., Liu S.-M., Chen Y.-T., Chang S.-J., Chen H.-C., Chen S.-K., Chen T.-R., Chen Y.-L., Cheng C.-H., Chung C.-I., Han S.-Y., Hsiao S.-H., Heiung J.-N., Hsu C.-H., Huang J.-J., Kan P.-I., Lee M.-C., Leu H.-L., Li Y.-F., Lin S.-J., Lin Y.-C., Wu S.-W., Yu C.-Y., Yu S.-W., Wu H.-P., Shaw J.-F., Submitted (NOV-2004) to the EMBL/GenBank/DBJ databases.	408	Qy
Db	573	RPT; P550293; TER REGION: 1. DR SEQUENCE 483 AA; 54448 MW; C652ED87A1A44DD2 CRC64;	573	RPT; P550293; TER REGION: 1. DR SEQUENCE 483 AA; 54448 MW; C652ED87A1A44DD2 CRC64;	573	Db
Qy	408	Query Match 16.6%; Score 570.5; DB 2; Length 483; Best Local Similarity 34.8%; Prod. No. 2.6e-15; Indels 49; Gaps 12; Matches 145; Conservative 69; Mismatches 154; Index 154; DR PROSITE: PS550293; TPR REGION: 1. DR SEQUENCE 483 AA; 54448 MW; C652ED87A1A44DD2 CRC64;	408	Query Match 16.6%; Score 570.5; DB 2; Length 483; Best Local Similarity 34.8%; Prod. No. 2.6e-15; Indels 49; Gaps 12; Matches 145; Conservative 69; Mismatches 154; Index 154; DR PROSITE: PS550293; TPR REGION: 1. DR SEQUENCE 483 AA; 54448 MW; C652ED87A1A44DD2 CRC64;	408	Qy
Db	573	9 AIFIQKWRHQARREMR-RCNWQFQNLAYASEDDQAELEYKPFNDLIKPMQAGRKN 67 77 ARYSKYTYRGRAYALMGKPIKFLAKQFQVERISPIDP-----DATRKKEFKEAVO 128 68 QYQGSIAHSVSLADDKODLVEERGDIIVAKIEL-----PIRKHIDLIDVFRKKRG 117 129 KIRPERAIISVGDEEKSVADSIDYR1IEVEFQYTGPRVGDGTIDLPVKAMIDDEFKQK- 187 118 NRLHPKVYVALLRANKSISQKLPNTIPVSTAVSQQTVCQDHLGKDDLIVLHNGGLPS 177 188 -CIIKR1AYQVLTQVOLRSPSLDQVNPDGSHP7VCGDVHGQFDLKLKFKLNGLPS 246	573	9 AIFIQKWRHQARREMR-RCNWQFQNLAYASEDDQAELEYKPFNDLIKPMQAGRKN 67 77 ARYSKYTYRGRAYALMGKPIKFLAKQFQVERISPIDP-----DATRKKEFKEAVO 128 68 QYQGSIAHSVSLADDKODLVEERGDIIVAKIEL-----PIRKHIDLIDVFRKKRG 117 129 KIRPERAIISVGDEEKSVADSIDYR1IEVEFQYTGPRVGDGTIDLPVKAMIDDEFKQK- 187 118 NRLHPKVYVALLRANKSISQKLPNTIPVSTAVSQQTVCQDHLGKDDLIVLHNGGLPS 177 188 -CIIKR1AYQVLTQVOLRSPSLDQVNPDGSHP7VCGDVHGQFDLKLKFKLNGLPS 246	573	Db
Qy	408	RESULTS 34 Q5W6J3 ORYSA PRELIMINARY; PRT; 483 AA.	408	RESULTS 34 Q5W6J3 ORYSA PRELIMINARY; PRT; 483 AA.	408	Qy
Db	573	Q5W6J3; 01-FEB-2005 (TREMBLrel. 29, Created) 01-FEB-2005 (TREMBLrel. 29, Last sequence update) 01-FEB-2005 (TREMBLrel. 29, Last annotation update)	573	Q5W6J3; 01-FEB-2005 (TREMBLrel. 29, Created) 01-FEB-2005 (TREMBLrel. 29, Last sequence update) 01-FEB-2005 (TREMBLrel. 29, Last annotation update)	573	Db
Qy	408	Name=OSJNB0115P21; 05W6J3;	408	Name=OSJNB0115P21; 05W6J3;	408	Qy
Db	573	Oryza sativa (Iaponica cultivar-group). Bukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzeae; Oryza. NCBI_TaxID=39947; [1]	573	Oryza sativa (Iaponica cultivar-group). Bukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzeae; Oryza. NCBI_TaxID=39947; [1]	573	Db
Qy	408	NUCLEOTIDE SEQUENCE. Chow T.-Y., Hsing Y.-I.C., Chen C.-S., Chen H.-H., Liu S.-M., Chen Y.-T., Chang S.-J., Chen H.-C., Chen S.-K., Chen T.-R., Chen Y.-L., Cheng C.-H., Chung C.-I., Han S.-Y., Hsiao S.-H., Heiung J.-N., Hsu C.-H., Huang J.-J., Kan P.-I., Lee M.-C., Leu H.-L., Li Y.-F., Lin S.-J., Lin Y.-C., Wu S.-W., Yu C.-Y., Yu S.-W., Wu H.-P., Shaw J.-F., Submitted (NOV-2004) to the EMBL/GenBank/DBJ databases.	408	NUCLEOTIDE SEQUENCE. Chow T.-Y., Hsing Y.-I.C., Chen C.-S., Chen H.-H., Liu S.-M., Chen Y.-T., Chang S.-J., Chen H.-C., Chen S.-K., Chen T.-R., Chen Y.-L., Cheng C.-H., Chung C.-I., Han S.-Y., Hsiao S.-H., Heiung J.-N., Hsu C.-H., Huang J.-J., Kan P.-I., Lee M.-C., Leu H.-L., Li Y.-F., Lin S.-J., Lin Y.-C., Wu S.-W., Yu C.-Y., Yu S.-W., Wu H.-P., Shaw J.-F., Submitted (NOV-2004) to the EMBL/GenBank/DBJ databases.	408	Qy
Db	573	RPT; P550293; TER REGION: 1. DR SEQUENCE 483 AA; 54448 MW; C652ED87A1A44DD2 CRC64;	573	RPT; P550293; TER REGION: 1. DR SEQUENCE 483 AA; 54448 MW; C652ED87A1A44DD2 CRC64;	573	Db
Qy	408	Putative serine/threonine phosphatase. Name=OSJNB0115P21; 05W6J3;	408	Putative serine/threonine phosphatase. Name=OSJNB0115P21; 05W6J3;	408	Qy
Db	573	Oryza sativa (Iaponica cultivar-group). Bukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzeae; Oryza. NCBI_TaxID=39947; [1]	573	Oryza sativa (Iaponica cultivar-group). Bukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzeae; Oryza. NCBI_TaxID=39947; [1]	573	Db
Qy	408	RESULTS 35 Q5W6J4 LYCES PRELIMINARY; PRT; 485 AA.	408	RESULTS 35 Q5W6J4 LYCES PRELIMINARY; PRT; 485 AA.	408	Qy
Db	573	Q5W6J4; 01-MAR-2003 (TREMBLrel. 23, Created) 01-MAR-2003 (TREMBLrel. 23, Last sequence update) 01-MAR-2003 (TREMBLrel. 23, Last annotation update)	573	Q5W6J4; 01-MAR-2003 (TREMBLrel. 23, Created) 01-MAR-2003 (TREMBLrel. 23, Last sequence update) 01-MAR-2003 (TREMBLrel. 23, Last annotation update)	573	Db
Qy	408	Type 5 serine/threonine phosphatase 55 kDa isoform (Type 5 protein berine/threonine phosphatase 55 kDa isoform). Name=PP5;	408	Type 5 serine/threonine phosphatase 55 kDa isoform (Type 5 protein berine/threonine phosphatase 55 kDa isoform). Name=PP5;	408	Qy
Db	573	LYCOPERSICON esculentum (Tomato). Bukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; lamials; Solanales; Solanaceae; Solanum; Lycopersicon. NCBI_TaxID=4081; [1]	573	LYCOPERSICON esculentum (Tomato). Bukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; lamials; Solanales; Solanaceae; Solanum; Lycopersicon. NCBI_TaxID=4081; [1]	573	Db
Qy	408	NUCLEOTIDE SEQUENCE. De la Fuente Van Bentem S., Vosken J.H., Vermeer J.B.M., Ra De Vroomen M.J., Gadella T.W.J., Haring M.A., Cornelissen B.J.; RT "The Subcellular localization of Plant Protein Phosphatase 5 isoforms is Determined by Alternative Splicing"; Plant Physiol. 133:702-712(2003).	408	NUCLEOTIDE SEQUENCE. de la Fuente Van Bentem S., Vosken J.H., Vermeer J.B.M., Ra De Vroomen M.J., Gadella T.W.J., Haring M.A., Cornelissen B.J.; RT "The Subcellular localization of Plant Protein Phosphatase 5 isoforms is Determined by Alternative Splicing"; Plant Physiol. 133:702-712(2003).	408	Qy
Db	573	RPT; P550293; TER REGION: 1. DR SEQUENCE 483 AA; 54448 MW; C652ED87A1A44DD2 CRC64;	573	RPT; P550293; TER REGION: 1. DR SEQUENCE 483 AA; 54448 MW; C652ED87A1A44DD2 CRC64;	573	Db
Qy	408	NUCLEOTIDE SEQUENCE. die la Fuente van Bentem S., Vosken J.H., Vermeer J.B.M., Ra De Vroomen M.J., Gadella T.W.J., Haring M.A., Cornelissen B.J.; RT "The Subcellular localization of Plant Protein Phosphatase 5 isoforms is Determined by Alternative Splicing"; Plant Physiol. 133:702-712(2003).	408	NUCLEOTIDE SEQUENCE. die la Fuente van Bentem S., Vosken J.H., Vermeer J.B.M., Ra De Vroomen M.J., Gadella T.W.J., Haring M.A., Cornelissen B.J.; RT "The Subcellular localization of Plant Protein Phosphatase 5 isoforms is Determined by Alternative Splicing"; Plant Physiol. 133:702-712(2003).	408	Qy
Db	573	RPT; P550293; TER REGION: 1. DR SEQUENCE 483 AA; 54448 MW; C652ED87A1A44DD2 CRC64;	573	RPT; P550293; TER REGION: 1. DR SEQUENCE 483 AA; 54448 MW; C652ED87A1A44DD2 CRC64;	573	Db
Qy	408	NUCLEOTIDE SEQUENCE. GO; GO:000573; C:cytoplasm; IEA.	408	NUCLEOTIDE SEQUENCE. GO; GO:000573; C:cytoplasm; IEA.	408	Qy
Db	573	INTERPRO: IPR006186; TPR; 3.	573	INTERPRO: IPR006186; TPR; 3.	573	Db
Qy	408	PROSITE; PS55005; TPR; 3.	408	PROSITE; PS55005; TPR; 3.	408	Qy
Db	573	PROSITE; PS550293; TER REGION: 1. DR SEQUENCE 483 AA; 54448 MW; C652ED87A1A44DD2 CRC64;	573	PROSITE; PS550293; TER REGION: 1. DR SEQUENCE 483 AA; 54448 MW; C652ED87A1A44DD2 CRC64;	573	Db
Qy	408	RESULTS 35 Q5W6J4 LYCES PRELIMINARY; PRT; 485 AA.	408	RESULTS 35 Q5W6J4 LYCES PRELIMINARY; PRT; 485 AA.	408	Qy
Db	573	Q5W6J4; 01-MAR-2003 (TREMBLrel. 23, Created) 01-MAR-2003 (TREMBLrel. 23, Last sequence update) 01-MAR-2003 (TREMBLrel. 23, Last annotation update)	573	Q5W6J4; 01-MAR-2003 (TREMBLrel. 23, Created) 01-MAR-2003 (TREMBLrel. 23, Last sequence update) 01-MAR-2003 (TREMBLrel. 23, Last annotation update)	573	Db
Qy	408	Type 5 serine/threonine phosphatase 55 kDa isoform (Type 5 protein berine/threonine phosphatase 55 kDa isoform). Name=PP5;	408	Type 5 serine/threonine phosphatase 55 kDa isoform (Type 5 protein berine/threonine phosphatase 55 kDa isoform). Name=PP5;	408	Qy
Db	573	LYCOPERSICON esculentum (Tomato). Bukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; lamials; Solanales; Solanaceae; Solanum; Lycopersicon. NCBI_TaxID=4081; [1]	573	LYCOPERSICON esculentum (Tomato). Bukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; lamials; Solanales; Solanaceae; Solanum; Lycopersicon. NCBI_TaxID=4081; [1]	573	Db
Qy	408	NUCLEOTIDE SEQUENCE. De la Fuente Van Bentem S., Vosken J.H., Vermeer J.B.M., Ra De Vroomen M.J., Gadella T.W.J., Haring M.A., Cornelissen B.J.; RT "The Subcellular localization of Plant Protein Phosphatase 5 isoforms is Determined by Alternative Splicing"; Plant Physiol. 133:702-712(2003).	408	NUCLEOTIDE SEQUENCE. de la Fuente Van Bentem S., Vosken J.H., Vermeer J.B.M., Ra De Vroomen M.J., Gadella T.W.J., Haring M.A., Cornelissen B.J.; RT "The Subcellular localization of Plant Protein Phosphatase 5 isoforms is Determined by Alternative Splicing"; Plant Physiol. 133:702-712(2003).	408	Qy
Db	573	RPT; P550293; TER REGION: 1. DR SEQUENCE 483 AA; 54448 MW; C652ED87A1A44DD2 CRC64;	573	RPT; P550293; TER REGION: 1. DR SEQUENCE 483 AA; 54448 MW; C652ED87A1A44DD2 CRC64;	573	Db
Qy	408	NUCLEOTIDE SEQUENCE. GO; GO:000573; C:cytoplasm; IEA.	408	NUCLEOTIDE SEQUENCE. GO; GO:000573; C:cytoplasm; IEA.	408	Qy
Db	573	INTERPRO: IPR006186; TPR; 3.	573	INTERPRO: IPR006186; TPR; 3.	573	Db
Qy	408	PROSITE; PS55005; TPR; 3.	408	PROSITE; PS55005; TPR; 3.	408	Qy
Db	573	PROSITE; PS55005; TPR; 3.	573	PROSITE; PS55005; TPR; 3.	573	Db
Qy	408	PROSITE; PS550293; TER REGION: 1. DR SEQUENCE 483 AA; 54448 MW; C652ED87A1A44DD2 CRC64;	408	PROSITE; PS550293; TER REGION: 1. DR SEQUENCE 483 AA; 54448 MW; C652ED87A1A44DD2 CRC64;	408	Qy
Db	573	INTERPRO: IPR01440; TPR; 3.	573	INTERPRO: IPR01440; TPR; 3.	573	Db
Qy	408	INTERPRO: IPR01190; TPR-like_helical.	408	INTERPRO: IPR01190; TPR-like_helical.	408	Qy
Db	573	INTERPRO: IPR006186; TPR; 3.	573	INTERPRO: IPR006186; TPR; 3.	573	Db
Qy	408	SMART; SM00156; PP2Ac; 1.	408	SMART; SM00156; PP2Ac; 1.	408	Qy
Db	573	PFAM; PF00515; TPR; 3.	573	PFAM; PF00515; TPR; 3.	573	Db
Qy	408	INTERPRO: IPR04843; M-peptidase.	408	INTERPRO: IPR04843; M-peptidase.	408	Qy
Db	573	PRINTS; PR0014; STPHPHASE.	573	PRINTS; PR0014; STPHPHASE.	573	Db
Qy	408	INTERPRO: IPR007242; P-intracellular signaling cascade; IEA.	408	INTERPRO: IPR007242; P-intracellular signaling cascade; IEA.	408	Qy
Db	573	DR GO:0006470; P:protein amino acid dephosphorylation; IEA.	573	DR GO:0006470; P:protein amino acid dephosphorylation; IEA.	573	Db
Qy	408	INTERPRO: IPR000909; PI_PLC_X.	408	INTERPRO: IPR000909; PI_PLC_X.	408	Qy
Db	573	INTERPRO: IPR01236; PP2Case_5.	573	INTERPRO: IPR01236; PP2Case_5.	573	Db
Qy	408	INTERPRO: IPR01440; TPR.	408	INTERPRO: IPR01440; TPR.	408	Qy
Db	573	INTERPRO: IPR01190; TPR-like_helical.	573	INTERPRO: IPR01190; TPR-like_helical.	573	Db
Qy	408	INTERPRO: IPR006186; TPR; 3.	408	INTERPRO: IPR006186; TPR; 3.	408	Qy
Db	573	SMART; SM00156; PP2Ac; 1.	573	SMART; SM00156; PP2Ac; 1.	573	Db
Qy	408	PROSITE; PS55007; PIPLC_X_DOMAIN; 1.	408	PROSITE; PS55007; PIPLC_X_DOMAIN; 1.	408	Qy
Db	573	PROSITE; PS55005; TPR; 3.	573	PROSITE; PS55005; TPR; 3.	573	Db
Qy	408	PROSITE; PS550293; TER REGION: 1. DR SEQUENCE 483 AA; 54448 MW; C652ED87A1A44DD2 CRC64;	408	PROSITE; PS550293; TER REGION: 1. DR SEQUENCE 483 AA; 54448 MW; C652ED87A1A44DD2 CRC64;	408	Qy
Db	573	INTERPRO: IPR01440; TPR; 3.	573	INTERPRO: IPR01440; TPR; 3.	573	Db
Qy	408	INTERPRO: IPR01190; TPR-like_helical.	408	INTERPRO: IPR01190; TPR-like_helical.	408	Qy
Db	573	INTERPRO: IPR006186; TPR; 3.	573	INTERPRO: IPR006186; TPR; 3.	573	Db
Qy	408	SMART; SM00156; PP2Ac; 1.	408	SMART; SM00156; PP2Ac; 1.	408	Qy
Db	573	PFAM; PF00515; TPR; 3.	573	PFAM; PF00515; TPR; 3.	573	Db
Qy	408	INTERPRO: IPR04843; M-peptidase.	408	INTERPRO: IPR04843; M-peptidase.	408	Qy
Db	573	PRINTS; PR0014; STPHPHASE.	573	PRINTS; PR0014; STPHPHASE.	573	Db
Qy	408	INTERPRO: IPR007242; P-intracellular signaling cascade; IEA.	408	INTERPRO: IPR007242; P-intracellular signaling cascade; IEA.	408	Qy
Db	573	DR GO:0006470; P:protein amino acid dephosphorylation; IEA.	573	DR GO:0006470; P:protein amino acid dephosphorylation; IEA.	573	Db
Qy	408	INTERPRO: IPR000909; PI_PLC_X.	408	INTERPRO: IPR000909; PI_PLC_X.	408	Qy
Db	573	INTERPRO: IPR01236; PP2Case_5.	573	INTERPRO: IPR01236; PP2Case_5.	573	Db
Qy	408	INTERPRO: IPR01440; TPR.	408	INTERPRO: IPR01440; TPR.	408	Qy
Db	573	INTERPRO: IPR01190; TPR-like_helical.	573	INTERPRO: IPR01190; TPR-like_helical.	573	Db
Qy	408	INTERPRO: IPR006186; TPR; 3.	408	INTERPRO: IPR006186; TPR; 3.	408	Qy
Db	573	SMART; SM00156; PP2Ac; 1.	573	SMART; SM00156; PP2Ac; 1.	573	Db
Qy	408	PROSITE; PS55007; PIPLC_X_DOMAIN; 1.	408	PROSITE; PS55007; PIPLC_X_DOMAIN; 1.	408	Qy
Db	573	PROSITE; PS55005; TPR; 3.	573	PROSITE; PS55005; TPR; 3.	573	Db
Qy	408	PROSITE; PS550293; TER REGION: 1. DR SEQUENCE 483 AA; 54448 MW; C652ED87A1A44DD2 CRC64;	408	PROSITE; PS550293; TER REGION: 1. DR SEQUENCE 483 AA; 54448 MW; C652ED87A1A44DD2 CRC64;	408	Qy
Db	573	INTERPRO: IPR01440; TPR; 3.	573	INTERPRO: IPR01440; TPR; 3.	573	Db
Qy	408	INTERPRO: IPR01190; TPR-like_helical.	408	INTERPRO: IPR01190; TPR-like_helical.	408	Qy
Db	573	INTERPRO: IPR006186; TPR; 3.	573	INTERPRO: IPR006186; TPR; 3.	573	Db
Qy	408	SMART; SM00156; PP2Ac; 1.	408	SMART; SM00156; PP2Ac; 1.	408	Qy
Db	573	PFAM; PF00515; TPR; 3.	573	PFAM; PF00515; TPR; 3.	573	Db
Qy	408	INTERPRO: IPR04843; M-peptidase.	408	INTERPRO: IPR04843; M-peptidase.	408	Qy
Db	573	PRINTS; PR0014; STPHPHASE.	573	PRINTS; PR0014; STPHPHASE.	573	Db
Qy	408	INTERPRO: IPR007242; P-intracellular signaling cascade; IEA.	408	INTERPRO: IPR007242; P-intracellular signaling cascade; IEA.	408	Qy
Db	573	DR GO:0006470; P:protein amino acid dephosphorylation; IEA.	573	DR GO:0006470; P:protein amino acid dephosphorylation; IEA.	573	Db
Qy	408	INTERPRO: IPR000909; PI_PLC_X.	408	INTERPRO: IPR000909; PI_PLC_X.	408	Qy
Db	573	INTERPRO: IPR01236; PP2Case_5.	573	INTERPRO: IPR01236; PP2Case_5.	573	Db
Qy	408	INTERPRO: IPR01440; TPR.	408	INTERPRO: IPR01440; TPR.	408	Qy
Db	573	INTERPRO: IPR01190				

SQ	SEQUENCE	485 AA;	54720 MW;	FCF68708D5520E68 CRC64;	[2]		
	Query Match	16.6%;	Score 570;	DB 2;	Length 485;		
	Best Local Similarity	36.1%;	Pred. No.	2.8e-26;			
	Matches	140;	Conservative	55;	Mismatches 124;		
					Indels 69; Gaps 12;		
QY	24	EMORRCNQIIFQNLLEYASBODQABLYKFFNDLTKHMPQAAGRKNQYQSAHVSVLDDKDD	83	RN	RN		
		: ::		RP	NUCLEOTIDE SEQUENCE.		
	142	ESQR--SVDSTDIYS-----	169	RC	NUCLEOTIDE-MUSCLE;		
Db				RA	STRAUERBERG R.;		
				RL	Submitted (NOV-2-2000) to the EMBL/GenBank/DDJB databases.		
QY	84	LVERFGDINAKTEPIRKRNHIDLLIDYFRKGRGNRILPKVALIRBAKSILKOLPNIS	143	RN	[3]		
		:   :		CC	NUCLEOTIDE SEQUENCE.		
	170	----GDVTL-----DFVKCMLDDFKNOK--NLKRYAYQVAF-----	214	CC	TISSUE-Muscle;		
Db				RA	Strauerberg R.;		
				RL	Submitted (JAN-2-2001) to the EMBL/GenBank/DDJB databases.		
QY	144	PVSTAVSQVTGCGDLHKGDLLWLNKGULPSSNPYVFFDVRGKLGIVLILL	203	RN	- I - CATALYTIC ACTIVITY: A phosphoprotein +		
		:     :		CC	PHOSPHATE.		
	215	DIVVPEGGKHFTVCGDVHQFYDPLNTIFLNGLPSEDNPYLFENGDFVDRGSLSLEVLTLP	274	CC	PHOSPHATE. Belongs to the PPP phosphatase family.		
Db				RA	- I - SIMILARITY: Belongs to the PPP phosphatase family.		
				RL	EMBL; BC000750 AAH00750 4; - mRNA.		
QY	204	SLVIAAPPNAFLNRGHNDSVMDARYGIRETESKYPAHNRKLAIFDEVVRLPLGSVL	263	RN	EMBL; BC001831 AAH01831.4; - mRNA.		
		:   :		CC	HSSP; P5301; 1A17.		
	275	AFKMCBCSAIHLARGNHESKSMKIKYGRGGRSKLRFVSEFA--EVFCCLPLAHV	331	CC	SMR; Q9BPW0; 21-481.		
Db				RA	GO; 0005737; Cytoplasm; IEA.		
				RL	GO; 0005634; C:nucleus; IEA.		
QY	264	NSRVLIVAGG-FS-DSTSLLDKLKSIDRKYKVSTLRPPLTDGEPLDKTETWQOQFDIMSDP	321	RN	DR; GO; 0005488; F:binding; IEA.		
		:   :		CC	DR; GO; 0016787; hydrolyase activity; IEA.		
	332	NEKVYVIGGLFSVDGVQLSDTAIDR---FCRPP-----BGMCELMISDP	376	CC	DR; GO; 0004721; F:phosphoprotein phosphatase activity; IEA.		
Db				RA	DR; GO; 0006470; P:protein amino acid dephosphorylation; IEA.		
				RL	DR; IPR004843; Metallophosphatase.		
QY	322	QATMGCVNTLRRAGWGVGPDTNPFORRALSYVTSHECKPGRHFMHNKNTITFSA	381	RN	DR; IPR011236; PPase_5.		
		:   :		CC	DR; IPR006186; T_Ptthase_apah.		
Db	377	QOPGRGPKSK-RGVLSPFQGDVTRFLQENNDLVRSHEVQDGYSIEHDQKLITYFSA	435	CC	DR; IPR011440; TPR.		
QY	382	SNTYAIGSNKGAYIRLN-NQLMHPHQY	408	CC	DR; IPR011990; TPR-like helical.		
		:   :		RA	DR; Pfam; PF00149; Metallophosphatase.		
Db	436	PNYCDOMANGKAGAIPFEEAPDMKPNIVTF	463	CC	DR; Pfam; PF00151; TPR_1.		
				RA	PIRSF; PIISF03096; PP2abe_5; 1.		
				CC	DR; PRINTS; PRO0114; STPPIIPTASE.		
				RA	DR; PRODom; P000252; T_phospha_apah; 1.		
				CC	DR; SMART; SW00156; PP2Ac; 1.		
				RA	DR; SMART; SW00028; TPR_1.		
				CC	DR; PROSTE; PS00125; SRR_TPR_PHOSPHATASE; UNKNOWN_1.		
				RA	DR; PROSTE; PS50005; TPR_3.		
				CC	DR; KW_Hydrolyse; IRON; TPR_REPEAT.		
				RA	FT_Non_TER_1		
				CC	SOQUENCE 484 AA; 55096 MW; 152A9C374E3D4F69 CRC64;		
				RA	Query Match 16.6%; Score 570; DB 2; Length 484;		
				CC	Best Local Similarity 33.7%; Pred. No. 2.8e-26;		
				RA	Matches 143; Conservative 71; Mismatches 142; Indels 68; Gaps 16;		
QY	11	FIOKWTRRHORREMORCRNQIIFQNLLEYASBODQABLYKFFNDLTKHMPQAAGRKNQY	70	CC	DB 136 ECNKTVKRAASNMALGK-----FR-----ALRD-----YETVVKVPHDKDAMKQ		
		:   :		RA	108 LIDVFRKKGRGILPKVALIRBAKSILKOLPNISPVNSTAVSQVTGCGDLHKGDLL 167		
	93	VIGGYIRRAASNMALGK-----FR-----ALRD-----YETVVKVPHDKDAMKQ	135	RA	192 LMQWYDOK--KLHRICKAYQIILWQVERKSLSTVTTKEBTKEITVCDPDTQGPYDLL 249		
QY	168	VVLRKGULPSSSNPYVFFNGDFDVRGKLGIVLILLISLYLAFFNAVFLNGRNHESVMA	227	RA	QY 168 VVLRKGULPSSSNPYVFFNGDFDVRGKLGIVLILLISLYLAFFNAVFLNGRNHESVMA 227		
		:   :		RA	DB 250 NIEPLKNGLPSPTNPYVFFNGDFDVRGKLGIVLILLISLYLAFFNAVFLNGRNHESVMA 309		
				RA	RA Rosak S.A., McBain P.J., McKernan K.J., Malek J.A., Gunaratne P.H., Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W., Villalon D.K., Muozz D.M., Soedergren E.J., Lu X., Gibbs R.A., Fahy J., Helton M., Madan A., Rodriguez S., Sanchez A., Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G., Blakesley R.W., Touchman J.W., Green R.D., Dickson M.C., Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.N., Krzywinski M.I., Skalska A., Schnarch A., Schein J.E., Jones S.J.M., Marra M.A., Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences.", Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).	RA	DB 228 RYGFPIREVESYPRKRILAFIDSYTMWPLGSVANSRVLIVHSGF--SDSTSLLIKS 285
QY	310	IYGEGEVKAKTY--AQMYELFSEVFWLPLAQCINGKVILMHGGLFSEPDGVTLDIRK	366	RA	DB 367 IENR-----OPP--DGPMP-----COLWSWPQONG-RS1SKRGYSCGPDDVTK 410		
QY	346	NFLQRHRLSYVTSHECKPNSHBFMMDKNTITFSSASNYAIGSNKGAYIRLN-NQLMHP	404	RA	DB 411 AFLEENLDVIRSHKABGIEVANGRCUTVFSAPNYCDOMGKASYIHLQGSDLRPQ 470		

QY	405	FVOY	408	RA	Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
ID	471	FHQF	474	RA	Bosak S.A., McElvan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
AC				RA	Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
ID	PF53041;	Q16722;		RA	Villalon D.K., Muzny D.M., Sodergren B.J., Lu X., Gibbs R.A.,
DT	01-OCT-1996	(Rel. 3.4, Created)		RA	Fahay J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
DT	01-OCT-1995	(Rel. 3.4, Last sequence update)		RA	Whiting M., Madan A., Young A.C., Shevchenko I., Bouffard G.G.,
DR	13-SEP-2005	(Rel. 48, Last annotation update)		RA	Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
DR	Serine/threonine protein phosphatase 5 (EC 3.1.3.16) (PP5) (protein phosphatase T) (PP-T) (PPT).			RA	Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
GN	Name=PP5C; Synonyms=PP5;			RA	Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.B.,
OS	Homo sapiens (Human);			RA	Schnetrich A., Schein J.E., Jones S.J.M., Marra M.A.;
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Buteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominoidea; Homo.			RT	"Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences;"
OC				RL	Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
NCBI_TAXID=9606;				RN	[5]
RN				RP	NUCLEOTIDE SEQUENCE OF 1-37.
RP				RX	Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RC				RA	NUCLEOTIDE SEQUENCE OF 7-499.
RX				RA	MEDLINE=5009929; PubMed=7925273;
RA				RA	Chen M.X., McPartlin A.E., Brown L., Chen Y.H., Barker H.M.,
RA	Cohen P.T.W.,			RA	Gubella J.P., Mohrenweiser H.W., Buckler A.J., Louis D.N.,
RT	"A novel human protein serine/threonine phosphatase, which possesses four tetratricopeptide repeat motifs and localizes to the nucleus.";			RT	"The structure of the tetratricopeptide repeats of the human protein kinase C-like protein phosphatase 5 gene.";
RT	EMBO J. 13:4278-4290 (1994).			RT	Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RL				RL	[1]
RN				RN	NUCLEOTIDE SEQUENCE OF 9-499.
RP				RN	MEDLINE=5009929; PubMed=7925273;
RC				RA	Xu X., Lagercrantz J., Zickert P., Bajajica-Lagercrantz S.,
RX				RA	Zetterberg A., Lagercrantz J., Zickert P., Bajajica-Lagercrantz S.,
RA	TISSUE=Fetal brain;			RA	Ollendorff V., Donoghue D.J.,
RA	Yong W.H., Ueki K., Chou D., Reeves S.A., von Deimling A.,			RA	"Chromosomal localization and 5' sequence of the human protein kinase C-like protein phosphatase 5 gene.";
RA	Gubella J.P., Mohrenweiser H.W., Buckler A.J., Louis D.N.,			RT	Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RT	"Cloning of a highly conserved human protein serine-threonine phosphatase gene from the glioma candidate region on chromosome 19q13.3.";			RT	Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RT	Genomics 29:533-536 (1995).			RT	[2]
RL				RL	NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RN				RN	MEDLINE=15057824; DOI=10.1088/nature03399;
RP				RP	GRIMWOOD J., Gordon L.A., Olsen A.S., Terry A., Schmitz J.,
RX				RP	Lamerdin J.B., Hallsten U., Goodstein D., Couronne O., Tran-Gyamfi M.,
RA	Arts A., Altherr M., Ashburner H., Bajorek E., Black S., Branccomb E.,			RP	"The structure of the tetratricopeptide repeats of protein phosphatase 5. Implications for TPR-mediated protein-protein interactions.";
RA	Caeneppele S., Carrasco A.V., Cauble C., Chan Y.M., Christensen M.,			RT	EMBO J. 17:1192-1199 (1998).
RA	Cleland C.A., Copeland A., Dalin B., Dehal P., Devys M., Detter J.C.,			RT	"The structure of the tetratricopeptide repeats of protein phosphatase 5. Implications for TPR-mediated protein-protein interactions.";
RA	Escobar J., Flores D., Fotopoulos D., Garcia C., Georgescu A.M.,			RT	EMBO J. 17:1192-1199 (1998).
RA	Glavina T., Gomez M., Gonzales E., Groza M., Hammom N., Hawkins T.,			RT	"The structure of the tetratricopeptide repeats of protein phosphatase 5. Implications for TPR-mediated protein-protein interactions.";
RA	Haydu L., Ho I., Huang W., Israely S., Jett J., Kadner K., Kimball H.,			RT	EMBO J. 17:1192-1199 (1998).
RA	Kobayashi A., Larionov V., Leem S.-H., Lopez F., Lou Y., Lowry S.,			CC	"The structure of the tetratricopeptide repeats of protein phosphatase 5. Implications for TPR-mediated protein-protein interactions.";
RA	Malفتati S., Martire D., McCready P.M., Medina C., Morgan J.,			CC	"The structure of the tetratricopeptide repeats of protein phosphatase 5. Implications for TPR-mediated protein-protein interactions.";
RA	Nelson K., Nolan M., Ovcharenko I., Pitluck S., Pollard M.,			CC	"The structure of the tetratricopeptide repeats of protein phosphatase 5. Implications for TPR-mediated protein-protein interactions.";
RA	Pedlik P., Prelik P., Quan G., Ramirez L., Rash S., Retterer J.,			CC	"The structure of the tetratricopeptide repeats of protein phosphatase 5. Implications for TPR-mediated protein-protein interactions.";
RA	Rodriguez A., Rogers S., Salamov A., Salazar A., She X., Smith D.,			CC	"The structure of the tetratricopeptide repeats of protein phosphatase 5. Implications for TPR-mediated protein-protein interactions.";
RA	Slezak T., Solovyev V., Thayer N., Tice H., Tsai M., Usashevskaya A.,			CC	"The structure of the tetratricopeptide repeats of protein phosphatase 5. Implications for TPR-mediated protein-protein interactions.";
RA	Vo N., Wagner M., Wheeler J., Wu K., Xie G., Yang J., Dubchak I.,			CC	"The structure of the tetratricopeptide repeats of protein phosphatase 5. Implications for TPR-mediated protein-protein interactions.";
RA	Forney T.S., DeJong P., Dickson D., Eichler E.B.,			CC	"The structure of the tetratricopeptide repeats of protein phosphatase 5. Implications for TPR-mediated protein-protein interactions.";
RA	Pennacchio L.A., Richardson P., Stubbs L., Rokhsar D.S., Myers R.M.,			CC	"The structure of the tetratricopeptide repeats of protein phosphatase 5. Implications for TPR-mediated protein-protein interactions.";
RA	Rubin E.M., Lucas S.M.,			CC	"The structure of the tetratricopeptide repeats of protein phosphatase 5. Implications for TPR-mediated protein-protein interactions.";
RA	"The DNA sequence and biology of human chromosome 19.";			CC	"The structure of the tetratricopeptide repeats of protein phosphatase 5. Implications for TPR-mediated protein-protein interactions.";
RA	Nature 428:529-535 (2004).			CC	"The structure of the tetratricopeptide repeats of protein phosphatase 5. Implications for TPR-mediated protein-protein interactions.";
RC	NUCLEOTIDE SEQUENCE [LARGE SCALE mRNA].			CC	"The structure of the tetratricopeptide repeats of protein phosphatase 5. Implications for TPR-mediated protein-protein interactions.";
RC	TISSUE=Cervix;			CC	"The structure of the tetratricopeptide repeats of protein phosphatase 5. Implications for TPR-mediated protein-protein interactions.";
RC	MEDLINE=2238257; PubMed=12477932; DOI=10.1073/pnas.242603899;			CC	"The structure of the tetratricopeptide repeats of protein phosphatase 5. Implications for TPR-mediated protein-protein interactions.";
RA	Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,			CC	"The structure of the tetratricopeptide repeats of protein phosphatase 5. Implications for TPR-mediated protein-protein interactions.";
RA	Klausner R.D., Collins F.S., Wagner L., Shemesh C.M., Schuler G.D.,			CC	"The structure of the tetratricopeptide repeats of protein phosphatase 5. Implications for TPR-mediated protein-protein interactions.";
RA	Achtsahl S.P., Zeeberg B., Butow K.H., Schaefer C.P., Bhat N.K.,			CC	"The structure of the tetratricopeptide repeats of protein phosphatase 5. Implications for TPR-mediated protein-protein interactions.";
RA	Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,			CC	"The structure of the tetratricopeptide repeats of protein phosphatase 5. Implications for TPR-mediated protein-protein interactions.";
RA	Diatchenko L., Matsunaga K., Farmer A.A., Rubin G.M., Hong L.,			CC	"The structure of the tetratricopeptide repeats of protein phosphatase 5. Implications for TPR-mediated protein-protein interactions.";
RA	Stapleton M., Soares M.P., Bonaldo M.P., Casavant T.L., Scheetz T.E.,			CC	"The structure of the tetratricopeptide repeats of protein phosphatase 5. Implications for TPR-mediated protein-protein interactions.";
RA	Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,			CC	"The structure of the tetratricopeptide repeats of protein phosphatase 5. Implications for TPR-mediated protein-protein interactions.";
RA	Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,			CC	"The structure of the tetratricopeptide repeats of protein phosphatase 5. Implications for TPR-mediated protein-protein interactions.";

DR GO; GO:004971; F:signal transducer activity; IMP.  
 DR GO; GO:007067; P:mitosis; TAS.  
 DR GO; GO:003123; P:protein regulation of I-kappaB kinase/NP-k- . . ; IMP.  
 DR GO; GO:006370; P:protein amino acid dephosphorylation; TAS.  
 DR GO; GO:006350; P:transcription; TAS.  
 DR InterPro; IPR01236; PPtase\_5.  
 DR InterPro; IPR006386; T\_Ptase\_apah.  
 DR InterPro; IPR01440; TPR.  
 DR InterPro; IPR011990; TPR-like\_helical.  
 DR Pfam; PF00149; Metallophos; 1.  
 DR Pfam; PF00515; TPR; 1; 3.  
 DR PIRSF; PIRSF033096; PPtase\_5; 1.  
 DR PRINTS; PR00114; STPHPTASE.  
 DR PRODOM; PD000252; T\_phosphatase\_apah; 1.  
 DR SMART; SM00156; PP2Ac; 1.  
 DR PROSITE; PS00125; SR\_THR\_PHOSPHATASE; 1.  
 DR PROSITE; PS00005; TPR; 3.  
 DR PROSITE; PS00293; TPR\_REGION; 1.  
 KW 3D-structure; Hydrolase; Iron; Manganese; Metal-binding;  
 Nuclear Protein; Protein phosphatase; Repeat; TPR repeat.  
 FT REPEAT 28 61  
 FT REPEAT 62 95  
 FT REPEAT 96 129  
 FT REGION 184 499  
 FT ACT SITE 304 304  
 FT METAL 242 242  
 FT METAL 244 244  
 FT METAL 271 271  
 FT METAL 303 303  
 FT METAL 352 352  
 FT METAL 427 427  
 FT CONFLICT 403 403  
 FT HELIX 22 40  
 FT TURN 41 42  
 FT HELIX 44 57  
 FT TURN 59 60  
 FT HELIX 62 74  
 FT TURN 75 76  
 FT HELIX 78 91  
 FT TURN 93 94  
 FT HELIX 96 108  
 FT TURN 109 110  
 FT HELIX 112 125  
 FT TURN 127 128  
 FT HELIX 130 164  
 FT TURN 168 169  
 SQ SEQUENCE 499 AA; 56879 MW; DB3B2090D8658BB3 CRC64;

Query Match 16.6%; Score 570; DB 1; Length 499;  
 Best Local Similarity 33.7%; Pred. No. 2.9e-25; Gaps 16;  
 Matches 143; Conservative 71; Mismatches 142; Indels 68; Gap 16;

OY 11 PIQKWMRHRQARERMRQRNQIFQNLEYASRHDQASLYKFENDLIGMPDAQRKQYQ 70  
 95 YIKGYRRRAANMAGK----FR---AALRD----YETVWVKPDKDARKYQ 137

OY 71 -----GSAHH-SVLDDKD---DLVSEFG--DIVAKIELEPIRKHIDL 107

Db 138 ECKKIVKQKAFFRALAGDEHKRSVVDSDIESMTIBYSGPKLEDGKVTSFPMKE--- 193

Db 108 LIDYFLVKRKGRLRHPKVYALLREBAAKSLQLPNISPVSTAVSQVTWCGDLAGKDDLL 167

Db 194 LMQWYDOKK--KLHKCAQYLQVQKVEVLKSTLVSTLKEETKTCGDIHQFYDLL 251

OY 168 WUHKGNGLPSSSPYVENGFDVDRGKQCLEVULLLSIYAPNAVTLRGHEDSYNA 227

252 NTFELNLGLPSETNPYIENGDFVDRGSPEVETLITLGPKLYDPDHFLRKGHETMNO 311

OY 228 RKGPIREVESKYPNRHKCILAFIDEVTRWLPLGSVLSNSRVLVHGGE--SISSTDLIKS 285

Db 312 IVGFEGEVKAKYT--AGMYELPSEVPFWLPLAQCINGKVLIMHGJGPSEBVTDIJK 368

RESULT 38  
 OY 053FR0\_HUMAN PRELIMINARY; PRT; 499 AA.  
 ID 053FR0\_HUMAN PRELIMINARY; PRT; 499 AA.  
 AC 053FR0/  
 DT 13-SEP-2005 (TMBimel. 31, Created)  
 DT 13-SEP-2005 (TMBimel. 31, Last sequence update)  
 DB 13-SEP-2005 (TMBimel. 31, Last annotation update)  
 OS Homo sapiens (human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi<sup>1</sup>;  
 OC Mammalia; Eutheria; Euarchontoglires; Primates; Cetartiodactyla;  
 OC Homo.  
 RN NCBI\_TAXID=9606;  
 [1]  
 RP NUCLEOTIDE SEQUENCE.  
 RC TISSUE=Stomach mucosa;  
 RA Maruyama K., Sugano S.;  
 RT "Oligo-capping: a simple method to replace the cap structure of  
 eucaryotic mRNAs with oligoribonucleotides.",  
 RL Gene 138:171-174(1994).  
 [2]  
 RN NUCLEOTIDE SEQUENCE.  
 RC TISSUE=Stomach mucosa;  
 RA Suzuki Y., Yoshimoto K., Maruyama K., Suyama A., Sugano S.;  
 RA Tanaka A., Tokoyama S.;  
 RL Submitted (APR-2005) to the EMBL/GenBank/DDBJ databases.  
 CC -I- CATALYTIC ACTIVITY: A phosphoprotein + H(2)O \* a protein +  
 phosphate. Belongs to the PPP phosphatase family.  
 CC -I- SIMILARITY: Belongs to the PPP phosphatase family.  
 DR EMBL; AK223222; BADB6942.1; --; mRNA.  
 DR SMR; 053FR0; 23-499.  
 DR GO; GO:0005737; C:cytoplasm; IEA.  
 DR GO; GO:0005634; C:nucleus; IEA.  
 DR GO; GO:0016787; F:hydrolase activity; IEA.  
 DR GO; GO:0004721; F:phosphoprotein phosphatase activity; IEA.  
 DR GO; GO:0006470; P:protein amino acid dephosphorylation; IEA.  
 DR InterPro; IPR004843; M-peptidase.  
 DR InterPro; IPR006186; T\_Ptase\_apah.  
 DR InterPro; IPR001440; TPR.  
 DR Pfam; PF00149; Metallophos; 1.  
 DR Pfam; PF00515; TPR; 3.  
 DR PRINTS; PR00114; STPHPTASE.  
 DR PRODOM; PD000222; T\_Ptase\_apah; 1.  
 DR SMART; SM00156; PP2Ac; 1.  
 DR SMART; SM00028; SR\_THR\_PHOSPHATASE; UNKNOWN\_1.  
 DR PROSITE; PS00125; SR\_THR\_PHOSPHATASE; UNKNOWN\_1.  
 DR PROSITE; PS50005; TPR; 3.  
 DR PROSITE; PS50293; TPR\_RKION; 1.  
 KW Hydrolase; Iron; TPR repeat.  
 FT NON\_TER 1 1  
 SQ SEQUENCE 499 AA; 56907 MW; DB5E20D85465C3F CRC64;

Query Match 16.6%; Score 570; DB 2; Length 499;  
 Best Local Similarity 33.7%; Pred. No. 2. 9e-26;  
 Matches 143; Conservative 71; Mismatches 142; Indels 68; Gaps 16;

DR Pfam; PF00149; Metallophos; 1.  
 DR Pfam; PF00515; TPR\_1; 3.  
 DR PRINTS; PRO0114; STPHPTASB.  
 DR PRODOM; PD000252; T\_Phtase\_apaH; 1.  
 DR SMART; SM00156; PP2AC; 1.  
 DR SMART; SM00228; TPR; 3.  
 DR PROSITE; PS00125; SER\_THR\_PHOSPHATASE; UNKNOWN\_1.  
 DR PROSITE; PS50005; TPR; 3.  
 DR PROSITE; PS50293; TPR\_REGION; 1.  
 DR Hydroxase; Iron; Repeat; TPR repeat.  
 KW Sequence; 499 AA; 56879 MW; DB3B090D08658BB3 CRC64;

Query Match 16.6%; Score 570; DB 2; Length 499;  
 Best Local Similarity 33.7%; Pred. No. 2. 9e-26;  
 Matches 143; Conservative 71; Mismatches 142; Indels 68; Gaps 16;

DR Pfam; PF00149; Metallophos; 1.  
 DR Pfam; PF00515; TPR\_1; 3.  
 DR PRINTS; PRO0114; STPHPTASB.  
 DR PRODOM; PD000252; T\_Phtase\_apaH; 1.  
 DR SMART; SM00156; PP2AC; 1.  
 DR SMART; SM00228; TPR; 3.  
 DR PROSITE; PS00125; SER\_THR\_PHOSPHATASE; UNKNOWN\_1.  
 DR PROSITE; PS50005; TPR; 3.  
 DR PROSITE; PS50293; TPR\_REGION; 1.  
 DR Hydroxase; Iron; Repeat; TPR repeat.  
 KW Sequence; 499 AA; 56879 MW; DB3B090D08658BB3 CRC64;

Query Match 16.6%; Score 570; DB 2; Length 499;  
 Best Local Similarity 33.7%; Pred. No. 2. 9e-26;  
 Matches 143; Conservative 71; Mismatches 142; Indels 68; Gaps 16;

DR Pfam; PF00149; Metallophos; 1.  
 DR Pfam; PF00515; TPR\_1; 3.  
 DR PRINTS; PRO0114; STPHPTASB.  
 DR PRODOM; PD000252; T\_Phtase\_apaH; 1.  
 DR SMART; SM00156; PP2AC; 1.  
 DR SMART; SM00228; TPR; 3.  
 DR PROSITE; PS00125; SER\_THR\_PHOSPHATASE; UNKNOWN\_1.  
 DR PROSITE; PS50005; TPR; 3.  
 DR PROSITE; PS50293; TPR\_REGION; 1.  
 DR Hydroxase; Iron; Repeat; TPR repeat.  
 KW Sequence; 499 AA; 56879 MW; DB3B090D08658BB3 CRC64;

RESULT 39

053XV2\_HUMAN

ID Q53XV2\_HUMAN PRELIMINARY; PRT; 499 AA.

AC 053XV2\_

DT 13-SEP-2005 (TREMBLrel. 31, Created)  
 13-SEP-2005 (TREMBLrel. 31, Last sequence update)  
 13-SEP-2005 (TREMBLrel. 31, Last annotation update)

DB protein\_in\_phosphatase\_5, catalytic\_subunit.

OS Homo Sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Buteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homindae; Homo.

NCBI\_TaxID=9606;

RN [1]

RR NUCLEOTIDE SEQUENCE.

RA Kainine N., Chen X., Rolfs A., Halleck A., Hines L., Eisenstein S., Phelan M., Farmer A., Raphael J., Moreira D., Kelley T., LaBaer J., Lin Y., "Cloning of human full-length cDNAs in BD Creator(TM) System Donor vector"; Submitted (MAY-2003) to the EMBL/GenBank/DDBJ databases.

RL -I CATALYTIC ACTIVITY: A phosphoprotein + H<sub>2</sub>O = a protein + phosphate.

CC -I SIMILARITY: Belongs to the PPP phosphatase family.

DR GO: GO:0016787; P:hydrolase activity; IFA.

DR EMBL; BT007275; AAP35939.1; -; mRNA.

DR Ensembl; ENSG0000011485; Homo sapiens.

DR GO; GO:005137; C:cycloplasm; IFA.

DR GO; GO:005634; C:nucleus; IFA.

DR GO; GO:0016787; P:hydrolytic enzyme activity; IFA.

DR GO; GO:0004721; P:protein phosphoprotein phosphatase activity; IFA.

DR GO; GO:0006470; P:protein amino acid dephosphorylation; IFA.

DR InterPro; IPR004843; M-pesterase.

DR InterPro; IPR006186; T\_phosphate\_aphA.

DR InterPro; IPR001440; TPR.

DR InterPro; IPR011990; TPR-like\_helical.

RESULT 40

064538\_RAT

ID Q64538\_RAT PRELIMINARY; PRT; 479 AA.

AC 064538\_

DT 01-NOV-1996 (TREMBLrel. 01, Created)  
 01-FEB-1997 (TREMBLrel. 02, Last sequence update)  
 01-MAR-2004 (TREMBLrel. 26, Last annotation update)

DB phosphoprotein\_Phosphatase (Fragment).

OS Rattus norvegicus (Rat).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Buteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

NCBI\_TaxID=10116;

RR NUCLEOTIDE SEQUENCE.

RC STRAITS-Sprague-Dawley; TISSUE-lung;

RX MEDLINE=95052208; PubMed=7972012;

RA Chinkers M.;

RT "Targeting of a distinctive protein-serine phosphatase to the protein kinase-like domain of the atrial natriuretic peptide receptor.";

RL Proc. Natl. Acad. Sci. U.S.A. 91:11075-11079 (1994).

CC -I CATALYTIC ACTIVITY: A phosphoprotein + H<sub>2</sub>O = a protein + phosphate.

-I SIMILARITY: Belongs to the PPP phosphatase family.

DR EMBL; U12203; AAB18614.1; - ; mRNA.  
 DR HSSP; P53041; 1A17.  
 DR SMR; 064538; 3-479.  
 DR GO; GO:000537; C:cytoplasm; IEA.  
 DR GO; GO:0005634; C:nucleus; IEA.  
 DR GO; GO:005488; F:binding; IEA.  
 DR GO; GO:001687; F:hydrolyase activity; IEA.  
 DR GO; GO:0004721; F:phosphoprotein phosphatase activity; IEA.  
 DR InterPro; IPR004843; M-peptidase.  
 DR InterPro; IPR006186; T\_phosphatase\_apah; 1.  
 DR InterPro; IPR01440; TPR.  
 DR InterPro; IPR01990; TPR-like helical.  
 DR Pfam; PF00149; Metallophos; 1.  
 DR Pfam; PF00515; TPR; 3.  
 DR PRINTS; PRO0114; STPIPIIHTASE.  
 DR PRODom; PD000252; T\_phosphatase\_apah; 1.  
 DR SMART; SM00155; PP2C; 1.  
 DR SMART; SM00028; TPR; 3.  
 DR PROSITE; PS00125; SER\_THR\_PHOSPHATASE; UNKNOWN\_1.  
 DR PROSITE; PS5005; TPR; 3.  
 DR PROSITE; PS50293; TPR\_REGION; 1.  
 DR KW Hydrolase; Iron; TPR repeat.  
 FT NON\_TPR 1  
 SQ 479 AA; 54718 MW; 173CAD1EFP526450 CIRG64;  
 Query Match 16.5%; Score 569; DB 2; Length 479;  
 Best Local Similarity 33.7%; Pred. No. 3.2e-26;  
 Matches 143; Conservative 73; Mismatches 140; Indels 68; Gaps 16;  
 QY 11 FLOKWYRRHQARREMRQCNWQTFONLYASRQDQABLYKFRNDLKKMPOAGRKNOQ 70  
 Db 75 YIKGYYRAASNNMALGK-----FR-----AALRD-----YETVVKUPDKADAKYQ 117  
 QY 71 -----GSAH-VSVDKDDK---DLVBERFG--DIVWAKIBPIRKHIDL 107  
 Db 118 ECSKIVKOKAFERRAIAGDEHRASVVDSDIDESMTIEDEYSGPRLKEDCKVTITFMKD---- 173  
 Qy 108 LIDVFERKRGNRHLHPKVVALIREAKSLQQLPNISVSTAVSQQVTCGDLHGKDDIL 167  
 Db 174 LMOWYKQOK-KLHRKCAYOIVNWKVKEVICKLSTLVETTLEKTEKLTIVCGDTHGQTYDLL 231  
 QY 168 VVUHKNGLPPSSNPYVNGDFUDRGKGELVWJULISLYLAEPNAVFLNRGHEHEDVMA 227  
 Db 232 NTFELNGLPSTEINPYFINGFDVDRGSFSEVILTCFPLKLYPDHFILRGHETDNMQ 291  
 QY 228 RYGFIREVESEKTPRNKHKLARLDEEVTRWLPLKGSVLNSRVLIVHGAF--SSTSLSLiks 285  
 Db 292 IVGPEGEVKAKT---AGWELRSEFVFWLPLAQCINGKVLMHGGLFSEDCVTLDDIRK 348  
 Qy 286 TIGKGVISLRLPPLTDERPLDKTNEWQQ1FDIIMWSDPQATMGCVPTLRCAGSMWFGDVTD 345  
 Db 349 IERNR-----QPP--DSGPW-----CDLWSDPQPNQRSVSK-RGVSCQFGDVTK 392  
 Qy 346 NFVQRHRISVIRSHECKPNQHBFMDNKITIFSAASNYAIGSNKGAYIRL-NNOLMPH 404  
 Db 393 AFLEENQDLYIIRSHEYGAEGEVAVHGCRGCVTVPSAPNQCOGMGNKASYIHLQSDLRQ 452  
 Qy 405 FVQY 408  
 Db 453 FHOF 456